Calhoun: The NPS Institutional Archive

DSpace Repository

Theses and Dissertations | 1. Thesis and Dissertation Collection, all items

2011-12

Measuring what right looks like: a system in developing metrics for tactical level units

Fuhr, Daniel A.; Pham, Hieu T.

Monterey, California. Naval Postgraduate School

http://hdl.handle.net/10945/10604

Downloaded from NPS Archive: Calhoun

http://www.nps.edu/library

Calhoun is a project of the Dudley Knox Library at NPS, furthering the precepts and goals of open government and government transparency. All information contained herein has been approved for release by the NPS Public Affairs Officer.

Dudley Knox Library / Naval Postgraduate School
411 Dyer Road / 1 University Circle
Monterey, California USA 93943

NAVAL POSTGRADUATE SCHOOL

MONTEREY, CALIFORNIA

THESIS

MEASURING WHAT RIGHT LOOKS LIKE: A SYSTEM IN DEVELOPING METRICS FOR TACTICAL LEVEL UNITS

by

Daniel Fuhr
December 2011

Hieu Pham
June 2012

Thesis Advisor: Hy Rothstein
Second Reader: James Henry

Approved for public release; distribution is unlimited

THIS PAGE INTENTIONALLY LEFT BLANK

REPORT DOCUMENTATION PAGE		Form Approved OMB No. 0704-0188

Public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instruction, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden, to Washington headquarters Services, Directorate for Information Operations and Reports, 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302, and to the Office of Management and Budget, Paperwork Reduction Project (0704-0188) Washington DC 20503.

1. AGENCY USE ONLY *(Leave blank)*	2. REPORT DATE December 2011	3. REPORT TYPE AND DATES COVERED Master's Thesis

4. TITLE AND SUBTITLE Measuring What Right Looks Like: A System in Developing Metrics for Tactical Level Units	5. FUNDING NUMBERS
6. AUTHOR(S) Daniel Fuhr and Hieu Pham	
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) Naval Postgraduate School Monterey, CA 93943-5000	8. PERFORMING ORGANIZATION REPORT NUMBER
9. SPONSORING /MONITORING AGENCY NAME(S) AND ADDRESS(ES) N/A	10. SPONSORING/MONITORING AGENCY REPORT NUMBER

11. SUPPLEMENTARY NOTES The views expressed in this thesis are those of the author and do not reflect the official policy or position of the Department of Defense or the U.S. Government. IRB Protocol number ___N/A____.

12a. DISTRIBUTION / AVAILABILITY STATEMENT Approved for public release; distribution is unlimited	12b. DISTRIBUTION CODE

13. ABSTRACT (maximum 200 words)

Since the United States' involvement in Afghanistan in 2001 and Iraq in 2003, thousands of U.S. service members have been lost and millions of man-hours spent on patrols, cordon and searches, and killing or capturing high value targets (HVTs). Billions of dollars from Commander's Emergency Response Program (CERP) have been spent on humanitarian aid projects. Despite this investment, outcomes remain vague.

This thesis devises a system for employment by tactical units to develop metrics that determine outcomes in nation assistance. It begins by defining terms and models useful for metric development in nation assistance: Rational Actor Theory, Dr. McCormick's Diamond Model, The Logic Model, and Correlation versus Causation. The thesis then uses historical examples of metrics from Vietnam, Iraq, and Afghanistan. Next, data analysis of nation assistance operations is reviewed. Difficulties and shortcomings in these historical examples and methods are highlighted. Finally, the thesis covers the Failed State Index that forms the base of the system that develops metrics that determine outcomes. The Tactical Outcome Assessment, was developed by operationalizing the Failed State Index for use by tactical units. The Tactical Outcome Assessment is the system that tactical units can employ to develop metrics that determine outcomes in nation assistance.

14. SUBJECT TERMS Nation Assistance, Metrics, The Failed State Index, Rational Actor, Dr. McCormick's Diamond Model	15. NUMBER OF PAGES 103
	16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT	18. SECURITY CLASSIFICATION OF THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT
Unclassified	Unclassified	Unclassified	UU

NSN 7540-01-280-5500

Standard Form 298 (Rev. 8-98)
Prescribed by ANSI Std. Z39.18

THIS PAGE INTENTIONALLY LEFT BLANK

Approved for public release; distribution is unlimited

MEASURING WHAT RIGHT LOOKS LIKE: A SYSTEM IN DEVELOPING METRICS FOR TACTICAL LEVEL UNITS

Daniel A. Fuhr
Major, United States Army
B.S., University of Massachusetts, 1999

Hieu T. Pham
Major, United States Army
B.S., University of Kansas, 1998

Submitted in partial fulfillment of the
requirements for the degree of

MASTER OF SCIENCE IN DEFENSE ANALYSIS

from the

NAVAL POSTGRADUATE SCHOOL
December 2011

Author: Daniel Fuhr
Hieu Pham

Approved by: Hy Rothstein
Thesis Advisor

James Henry
Second Reader

John Arquilla
Chair, Department of Defense Analysis

THIS PAGE INTENTIONALLY LEFT BLANK

ABSTRACT

Since the United States' involvement in Afghanistan in 2001 and Iraq in 2003, thousands of U.S. service members have been lost and millions of man-hours spent on patrols, cordon and searches, and killing or capturing high value targets (HVTs). Billions of dollars from Commander's Emergency Response Program (CERP) have been spent on humanitarian aid projects. Despite this investment, outcomes remain vague.

This thesis devises a system for employment by tactical units to develop metrics that determine outcomes in nation assistance. It begins by defining terms and models useful for metric development in nation assistance: Rational Actor Theory, Dr. McCormick's Diamond Model, The Logic Model, and Correlation versus Causation. The thesis then uses historical examples of metrics from Vietnam, Iraq, and Afghanistan. Next, data analysis of nation assistance operations is reviewed. Difficulties and shortcomings in these historical examples and methods are highlighted. Finally, the thesis covers the Failed State Index that forms the base of the system that develops metrics that determine outcomes. The Tactical Outcome Assessment, was developed by operationalizing the Failed State Index for use by tactical units. The Tactical Outcome Assessment is the system that tactical units can employ to develop metrics that determine outcomes in nation assistance.

THIS PAGE INTENTIONALLY LEFT BLANK

TABLE OF CONTENTS

I. INTRODUCTION 1
A. SCOPE OF THESIS – A SYSTEM FOR DEVELOPING METRICS 2
B. METHODOLOGY – THE USE OF A CASE STUDY TO OPERATIONALIZE THE FAILED STATE INDEX 5
C. INTRODUCTION OF CHAPTERS 7
1. Chapter II – Background for Metrics 7
2. Chapter III – When in Doubt of What You Should Measure, Measure What You Can 7
4. Chapter V – Operationalizing the Failed State Index for Measuring Outcomes at the Tactical Level 8
5. Chapter VI – Conclusion and Recommendations 8

II. BACKGROUND FOR METRICS 9
A. THE RATIONAL ACTOR MODEL 9
B. DR. GORDON MCCORMICK'S DIAMOND MODEL 11
C. THE LOGIC MODEL 13
D. PERFECTION IS THE ENEMY OF GOOD ENOUGH: CORRELLATION VS. CAUSATION 15
1. Correlation 15
2. Causation 16
3. Experience and Correlation 16
4. Correlation Versus Causation in the Fields of Health and Medicine 17
5. Correlation Versus Causation and the Relationship of Dependent and Independent Variables 19

III. WHEN IN DOUBT OF WHAT YOU SHOULD MEASURE, MEASURE WHAT YOU CAN 21
A. INTRODUCTION 21
B. METRICS IN VIETNAM 23
1. Hearts and Minds and Body Count 23
2. The Problems with Body Count 25
3. Problems with Other Metrics in Vietnam 27
4. Conclusion 28
C. METRICS IN IRAQ AND AFGHANISTAN 29
1. Qualitative and Quantitative Metrics 29
a. Tactical Conflict Assessment and Planning Framework (TCAPF) 30
b. Measuring Progress in Conflict Environments (MPICE) 33
c. United States Agency for International Development's "F" Framework 34

d. RAND and Brookings Institute MOEs for Counterinsurgency 35
e. Polling Data 35
f. David Kilcullen's MOEs from Counterinsurgency.... 35
2. Conclusion 36

IV. DATA ANALYSIS – NUMBERS DO NOT LIE BUT THEY ARE OPEN TO INTERPRETATION 37
A. COMMONLY USED DATA REPOSITORIES 37
B. DIFFICULTY IN SELECTING A DEPENDENT VARIABLE IN NATION ASSISTANCE 38
1. Selection of the Dependent Variable 38
2. What is Representative of the Dependent Variable? 39
3. Historic Problems Representing the Dependent Variable in Data Analysis 40
5. A Viable Dependent Variable 41
C. DIFFICULTIES IN DATA ANALYSIS 41
1. Problems with Accuracy of Data Sets 41
2. No Smoking Gun – The Results of Data Analysis 43
3. In a One-Chance, Zero Sum Game Where Every Option Has a Cost, What Probability Do You Need for Action? 47

V. OPERATIONALIZING THE FAILED STATE INDEX FOR MEASURING OUTCOMES AT THE TACTICAL LEVEL 51
A. INTRODUCTION 51
B. NATION BUILDING 52
C. THE FAILED STATE INDEX 52
1. Introduction to the Failed State Index 52
2. The Twelve Indicators of the Failed State Index 54
3. Sources of Data for the Failed State Index 55
4. Scoring the Failed State Index 55
5. Uses for the Failed State Index 56
6. Characteristics of Failed States 57
7. Conclusion 58
D. SADR CITY CASE STUDY 58
1. Background and History of Sadr City 58
2. Nation Building at the Tactical Level 61
E. EVALUATING THE TACTICAL UNIT ON THE FAILED STATE INDEX 62
1. Operationalizing the Failed State Index 62
2. Evaluating a Tactical Unit in Sadr City with the Operationalized Failed State Index 66
3. Tactical Outcomes Nesting With the Broader Strategy in U.S. Nation Assistance Missions 69

VI. CONCLUSION AND RECOMMENDATIONS 71
A. CONCLUSION 71

1. Theories and Models Useful in Determining Metrics That Measure Outcomes 71
a. Rational Actor Model 72
b. Dr. Gordon McCormick's Diamond Model 72
c. The Logic Model 73
d. Correlation vs. Causation 74
2. What is Not Effective and What is Not Useful in Determining Metrics that Measure Outcomes 75
a. Overly Complex Qualitative and Overly Detailed Quantitative Metrics that Fail to Measure Outcomes 75
b. Data Analysis 75
3. Tactical Outcome Assessment 76
B. RECOMMENDATIONS 77
1. Implementation 77
2. Areas for Further Research 77
LIST OF REFERENCES 79
INITIAL DISTRIBUTION LIST 85

THIS PAGE INTENTIONALLY LEFT BLANK

LIST OF FIGURES

Figure 1. Dr. Gordon McCormick's Diamond Model 11
Figure 2. Example Logic Model 14
Figure 3. Sample TCAPF from Iraq 31
Figure 4. Kunar & Kabul Area Indicators of Popular Support, Causes of Instability, Friendly Actions and CERP Expenditure 44
Figure 5. Kandahar Indicators of Popular Support, Causes of Instability, Friendly Actions and CERP Expenditure 44
Figure 6. Helmand Indicators of Popular Support, Causes of Instability, Friendly Actions and CERP Expenditure 45
Figure 7. Unemployment Rate vs. Violence Rate in Salinas, CA..... 47
Figure 8. Scatter-plot of Violent Acts per CERP $ Spent in At Tameem..... 49
Figure 9. Failed State Indicators..... 54
Figure 10. Map of Sadr City..... 60
Figure 11. Tactical Outcome Assessment 63
Figure 12. Electrical Wires in Adhamiyah 64
Figure 13. Tactical Outcome Assessment Filled Out..... 68

THIS PAGE INTENTIONALLY LEFT BLANK

LIST OF ACRONYMS AND ABBREVIATIONS

AO	Area of Operation
BCT	Brigade Combat Team
CAST	Conflict Assessment System Tool
CCIR	Commander's Critical Information Requirements
CERP	Commander's Emergency Response Program
CIDNE	Combined Data Network Exchange
FHA	Foreign Humanitarian Assistance
FID	Foreign Internal Defense
GOI	Government of Iraq
HCA	Humanitarian and Civic Assistance
HN	Host Nation
HVT	High Value Target
IDP	Internally Displaced Persons
IED	Improvised Explosive Device
IPB	Intelligence Preparation of the Battlefield
ISAF	International Security Assistance Forces
JAM	Jay'sh al-Mahdi
JSOTF	Joint Special Operations Task Force
LLO	Logical Lines of Operations
MACV	Military Assistance Command Vietnam
MEDCAP	Medical Civic Action Program
MNC-I	Multi-National Corps – Iraq

MOE	Measure of Effectiveness
MPICE	Measuring Progress in Conflict Environments
NA	Nation Assistance
OE	Operating Environment
OIF	Operation Iraqi Freedom
PIR	Priority Intelligence Requirements
SA	Security Assistance
SIGACT	Significant Activity
SOTF	Special Operations Task Force
TCAPF	Tactical Conflict Assessment and Planning Framework
TIGR	Tactical Integrated Ground Reporting
TTP	Tactics, Techniques, and Procedures
UNDP	United Nations Development Program
UNHCR	United Nations High Commission of Refugees
VC	Viet Cong
WHO	World Health Organization
WITS	Worldwide Incidents Tracking System

ACKNOWLEDGMENTS

First we would like to thank our advisor, Dr. Hy Rothstein and second reader, Major James Henry for entertaining our attempt at developing a system for developing metrics in nation assistance operations. We would like thank Dr. Fox for his help and patience in explaining mathematical concepts to people who still rely on their fingers for adding and subtracting. We would like to thank the Defense Analysis Department for affording us this opportunity. Finally, we would like to thank our friends and families.

THIS PAGE INTENTIONALLY LEFT BLANK

I. INTRODUCTION

Knowing what to measure and how to measure it makes a complicated world much less so.[1]

In 2003, shortly after the invasion of Iraq, former Secretary of Defense Donald Rumsfeld wrote in a memorandum, "Today, we lack metrics to know if we are winning or losing the global war on terror."[2] Since the United States' involvement in Afghanistan in 2001 and Iraq in 2003, thousands of U.S. service members have been lost, millions of man-hours spent on patrols, cordon and searches, and killing or capturing high value targets (HVTs). Billions of dollars from Commander's Emergency Response Program (CERP) have been spent on humanitarian aid projects.[3] Despite this investment, significantly less has been done to determine what outcomes these actions achieved.

It is too common to hear service members return from recent deployments in Iraq and Afghanistan and say, "We don't know what effect we really had." One service member who returned from Operation Enduring Freedom – Philippines in 2009 told how the number of circumcisions performed during medical civic action programs (MEDCAP) was a metric being tracked, as if the number of circumcisions correlated to increased stability in the area.[4] Currently, tactical level measures of effectiveness (MOEs) do not tell us if we are winning or losing the situations in which the U.S. military is involved.

---

[1] Steven D. Levitt and Stephen J. Dubner, *Freakonomics: A Rogue Economist Explores the Hidden Side of Everything* (New York: Harper Collins, 2006, 14).

[2] Donald Rumsfeld, *Rumsfeld Memo*, October 13, 2003, http://www.usatoday.com/news/washington/executive/rumsfeld-memo.htm (accessed January 27, 2011).

[3] In a report from February 2010, the U.S. Congress has appropriated more than $53 billion for Iraq's reconstruction and $51 billion towards Afghanistan's reconstruction. Arnold Fields, "Testimony of Arnold Fields Special Inspector General for Afghanistan Reconstruction Before the Commission on Wartime Contacting," *SIGAR.mil*, February 22, 2010, http://www.sigar.mil/pdf/testimony/SIGAR-10-001T.pdf (accessed November 01, 2010).

[4] This information came from a peer discussion during a class at Naval Postgraduate School, Monterey, CA, March, 2011.

A. SCOPE OF THESIS – A SYSTEM FOR DEVELOPING METRICS

The purpose of this thesis is to answer two questions:

- What metrics should tactical level military units employ for measuring outcomes in nation assistance (NA) operations?
- How should units' measures of effectiveness in their specific operating environments connect to metrics of their command and ultimately nest with the broader strategy in U.S. NA operations?

These questions stem from the assumption that the current operations in Iraq and Afghanistan align more with the definition of NA than of COIN. To answer the aforementioned questions, a system for use by tactical units will be developed that, step-by-step, guides them in developing metrics for their operating environment (OE) that are (1) accurate, (2) within a unit's ability to collect, (3) informative, and which (4) determine where and how to spend time and money.[5] First, accurate is used to mean that there is correlation between the outcome (dependent variable) and the metric (independent variables). The metric must have relationship to the outcome. If a unit's goal is to protect the populace,[6] the metric used to measure progress towards protecting the populace must identify whether or not progress is being made. We hope but do not expect to find actual causation. We argue that correlation and experience is enough for action. This will be covered in detail in Chapter IID. Second, The metrics need to be within the unit's ability to collect. The complexity of nation assistance operations has had a tendency to spark ambition and create metrics beyond what a tactical unit can accurately report. Chapter III discusses the historical problems of over-measuring in nation assistance operations and measuring the wrong indicators. Metrics from Vietnam will be discussed in this chapter, as well

---

[5] MAJ Geoffrey Van Epps originated the idea that the goal of a metric is to determine where and how to spend time and money during OIF 09-10.

[6] Protecting the populace has been a recommended and stated goal for both Iraq and Afghanistan. Elliot Cohen, Conrad Crane, Jan Horvath and John Nagle, "Principles, Imperatives, and Paradoxes of Counterinsurgency," Military *Review*, March-April 2006: 52. General Petraeus reiterates the goal of protecting the populace while addressing NATO in July 2010 shortly after replacing General McChrystal as the new head of U.S. and NATO forces in Afghanistan, "You must do everything possible to protect the population." *US chief Petraeus vows to protect Afghan civilians*, July 1, 2010, http://www.bbc.co.uk/news/10472555 (accessed March 11, 2011).

as metrics from current operations in Iraq and Afghanistan. Third, the metric must be informative. The metric must be able to tell the unit whether or not they are moving closer or farther from their outcome. Fourth, the metric must inform the unit's decision-making loop in determining where and how to spend time and money. The metric must be actionable.

Unfortunately, commonly used metrics in today's operations in Iraq and Afghanistan revolve around the kill or capture of HVTs, the number and costs of development projects, or the populace's access to essential services.[7] The latter are rife with problems as they fail to measure outcomes.[8] For instance, if we examine the populace's access to essential services, the problems of this metric will be apparent. First, access to essential services may not relate to the goal of protecting the populous. Having or not having a modern water treatment facility may not affect stability in the area; there may be no correlation of a modern water treatment facility to stability, let alone causation. Second, a tactical unit does not have the capability to accurately measure access to essential services. A tactical unit does not have the time, expertise, or resources to conduct a census of the population in their OE and then conduct a survey with enough depth to accurately measure what percent has access. Third, an accurate measure of the populace's access to a modern water treatment facility does not assist the unit in telling whether they are moving farther or closer towards stabilization and

[7] The kill or capture of HVTs, number and amount of development projects, and access to essential services were metrics actively being tracked at Multinational Division Iraq during 2007-2010. Another commonly tracked metric was dollars spent. David Brooks wrote this in a *New York Times* article: "Many in Congress fixate on "burn rates" — how fast a program can disperse money — not effectiveness." This effect trickles down to the tactical level where money was used more as a metric system than as a weapons system. *Smart Power Setback*, June 20, 2011, http://www.nytimes.com/2011/06/21/opinion/21brooks.html (accessed July 07, 2011). Discussions with peers who have deployed to Iraq and Afghanistan as staff primaries at the JSOTF and BCT level confirm the above were actively tracked on their deployments.

[8] Stephen Downes-Martin discusses the flawed logic and inaccurate use of numbers in operation assessments in Afghanistan in his article: Stephen Downes-Martin, "Operations Assessment in Afghanistan is Broken: What is to be Done? [PDF]," *Naval War College Review*, Autumn 2011: 103–125.

security objectives.[9] Fourth, even if a modern water treatment facility was the key to protecting the populace, it would not be actionable. A tactical unit would not be able to build a multi-million-dollar water treatment facility requiring a multi-year build along with the necessary staff needed for running the facility. Therefore, knowing that a functioning water treatment facility is needed for stability would not be relevant for a unit in determining where and how to spend time and money.

However, the example of the water treatment facility requires a note. Tracking essential service projects does have merit as a metric, but not merit directly from the project itself. Tracking essential service projects and the rate of completion of essential service projects executed by the host nation government is a valid metric for measuring the functionality of the host nation government. A host nation government must have some form of local authority (with an actual budget) in a given area in order to execute an essential service project. A budget is representative of a legitimate revenue based on some level of communication between the government and the population. Essential service projects also show that the government must be functioning well enough. Essential service projects started and completed by the host nation government is a valid metric that shows government capability.

Essential service projects started and completed by the host nation government also build legitimacy for the host nation government. In counterinsurgency, and nation assistance, the populace and its belief in and support of its government is the center of gravity.[10] If the populace sees that its government is able to provide increasing access to essential services, the host nation government will gain legitimacy. This legitimacy will be undermined if external actors (U.S. Forces in the case of Iraq and Afghanistan) are seen as the

---

[9] "A conference of experts brought together last year in Wilton Park in Britain concluded that there is a 'surprisingly weak evidence base for the effectiveness of aid in promoting stabilization and security objectives' in Afghanistan." David Brooks, *Smart Power Setback*, June 20, 2011, http://www.nytimes.com/2011/06/21/opinion/21brooks.html (accessed July 07, 2011).

[10] Kalev I. Sepp, "Best Practices in Counterinsurgency," *Military Review*, May/June 2005: 9.

ones behind the project and not the host nation government. Unfortunately, in Iraq and Afghanistan, the projects themselves were the goal and not the legitimacy of the host nation government. This results in combat philanthropy or what Bernard Fall referred to when instructing on how Vietnam cannot be won by giving out better privies.

> One side says, "land reform," and the other side says, "better culverts." One side says, "We are going to kill all those nasty village chiefs and landlords." The other side says, "Yes, but look, we want to give you prize pigs to improve your strain."[11]

This strategy failed in Vietnam and is failing to produce tangible effects in Iraq and Afghanistan.[12]

Instead of measuring misadventures in philanthropy or counting the number of skins on the wall (kill/capture of HVTs), this thesis seeks to create a system that allows tactical units to measure what matters in order to facilitate political outcomes in their OE. The metrics must meet the four previous requirements. The capabilities of the average patrol operating in a conflict environment need to be taken into account so the metrics they are being asked to report are in line with the unit's capability to collect.

B. METHODOLOGY – THE USE OF A CASE STUDY TO OPERATIONALIZE THE FAILED STATE INDEX

This thesis seeks to answer the questions:

- What metrics should tactical-level military units employ for measuring outcomes in nation assistance (NA) operations?

---

[11] Bernard B. Fall, "The Theory and Practice of Insurgency and Counterinsurgency," *The Naval War College Review*, April 1965: 55.

[12] David Brooks, "Smart Power Setback," June 20, 2011, http://www.nytimes.com/2011/06/21/opinion/21brooks.html (accessed July 07, 2011). And, Edwina Thompson, *WINNING 'HEARTS AND MINDS' IN AFGHANISTAN: ASSESSING THE EFFECTIVENESS OF DEVELOPMENT AID IN COIN OPERATIONS*, April 2010, http://www.wiltonpark.org.uk/resources/en/pdf/22290903/22291297/wp1022-report (accessed July 23, 2011).

- How should a unit's measures of effectiveness in their specific operating environment connect to metrics of their command and ultimately nest with the broader strategy in U.S. NA operations?

The definition of nation assistance from Joint Publication 3-22, Foreign Internal Defense is:

> Civil or military assistance (other than foreign humanitarian assistance [FHA]) rendered to a nation by US forces within that nation's territory during peacetime, crises or emergencies, or war, based on agreements mutually concluded between the United States and that nation. NA operations support the host nation (HN) by promoting sustainable development and growth of responsive institutions. The goal is to promote long-term regional stability. NA programs include security assistance (SA), humanitarian and civic assistance (HCA), and foreign internal defense (FID). FID is the participation by civilian and military agencies of a government in any of the action programs taken by another government or other designated organization, to free and protect its society from subversion, lawlessness, insurgency, terrorism, and other threats to their security.[13]

Again, the goal of nation assistance is to promote long-term regional stability. In order for the latter two questions to be answered, the metrics for use in nation assistance must be capable of measuring progress towards or away from long-term regional stability. This thesis will utilize a case study of a tactical unit's operations inside Sadr City during Operation Iraqi Freedom 08-09.

Next, this thesis hypothesizes that the twelve indicators of the Failed State Index developed by the Fund for Peace,[14] operationalized for use by a tactical unit, are effective metrics in nation assistance. First, this case study will show that the unit's actions and individual missions inside Sadr City were nation assistance. Second, the unit's actions will be shown to have impact on the

---

[13] Joint Publication 3-22: Foreign Internal Defense (2010, I-1).

[14] "The Fund for Peace is an independent, nonpartisan, 501(c)(3) non-profit research and educational organization that works to prevent violent conflict and promote sustainable security." Every year this organization ranks 177 countries across twelve indicators of state effectiveness. This becomes the Failed State Index that is published annually by Foreign Policy magazine. *About the Fund for Peace*, 2011, http://www.fundforpeace.org/global/?q=aboutus (accessed August 21, 2011).

indicators from the Failed State Index. Third, the indicators from the Failed State Index, operationalized for use, will be used to evaluate the outcomes produced in Sadr City by the tactical unit.

C. INTRODUCTION OF CHAPTERS

This thesis will be composed of six chapters including the introduction.

1. Chapter II – Background for Metrics

This chapter will review basic models and definitions used in developing metrics for nation assistance operations. The rational actor model, used by Thomas Schelling,[15] will be applied to nation assistance in order to explain the motivation of the populace. This will lead into Dr. McCormick's Diamond Model.[16] The Diamond is a useful model in analyzing an area as it contains a feedback mechanism that can be used to explain increasing or decreasing success as the population support moves from the insurgent to the state or from the state to the insurgent. Next, the logic model will be used to explain the differences between inputs, throughputs, outputs, and outcomes.[17] These definitions are necessary for determining metrics that measure actual results and not performance results. Finally, the chapter will conclude with definitions and explanation of correlation and causation. Chapter V will utilize these models and definitions in developing metrics for nation assistance operations.

2. Chapter III – When in Doubt of What You Should Measure, Measure What You Can

This chapter begins with historical problems in metrics starting with body count in Vietnam. The chapter will continue to discuss the difficulties in developing metrics moving into the current conflicts in Iraq and Afghanistan.

---

[15] Thomas C. Schelling, *Arms and Influence* (New Haven, Connecticut: Yale University Press, 1966, 38–58).

[16] Dr. McCormick's model of insurgency is best captured in Eric P. Wendt, "Strategic Counterinsurgency Modeling," *Special Warfare*, September 2005: 2–13.

[17] "Logic Model Development Guide" [PDF] (W.K. Kellogg Foundation, 2004, 1).

3. Chapter IV – Data Analysis: Numbers Do Not Lie, But They Are Open to Interpretation

This chapter reviews our data analysis as well as our analysis of others' explanations of their data analysis in nation assistance operations. Prior to discussing our data and others' data, we discuss briefly the difficulty of selecting a viable dependent variable. Levels of violence and number of reported tips, commonly used dependent variables in data analysis, may not correlate to long-term stability. Long-term stability is a difficult concept and hard to capture as a variable. This will lead into the inconclusive results of past attempts at data analysis. Theses containing data analysis by David Beskow, Thomas O'Connell, Justin Gorkowski, and Jason Clark and Tracy Onufer will also be discussed.

4. Chapter V – Operationalizing the Failed State Index for Measuring Outcomes at the Tactical Level

This chapter will start with an explanation of nation building as this activity is directly related to addressing the Failed State Index indicators. Next, this chapter will outline the Failed State Index developed by The Fund for Peace. The twelve indicators in the index will be fully explained to show how the indicators are measured. Then, the chapter will show how a tactical unit's operations in Sadr City addressed many of the indicators in the Failed State Index. This chapter will conclude with the indicators from the Failed State Index being operationalized into the Tactical Outcome Assessment and used in the case study of Sadr City to measure the tactical unit's outcomes.

5. Chapter VI – Conclusion and Recommendations

This chapter summarizes the previous chapters. Models and definitions useful for the development of metrics and what are neither effective nor useful in determining metrics that measure outcomes in nation assistance will be reviewed. The system this thesis developed for measuring outcomes in nation assistance will then be summarized. The thesis will conclude with recommendations on implementation and areas that require further research.

II. BACKGROUND FOR METRICS

This chapter will review basic models and definitions used in developing metrics for nation assistance operations. The rational actor model, used by Thomas Schelling,[18] will be applied to nation assistance in order to explain the motivation of the populace. This will lead into Dr. McCormick's Diamond Model.[19] The Diamond is a useful model in analyzing an area as it contains a feedback mechanism that can be used to explain increasing or decreasing success as the population support moves from the insurgent to the state or from the state to the insurgent. Next, the logic model will be used to explain the differences between inputs, throughputs, outputs, and outcomes.[20] These definitions are necessary for determining metrics that measure actual results and not performance results. Finally, the chapter will conclude with definitions and explanation of correlation and causation. Chapter V will utilize these models and definitions in developing metrics for nation assistance operations.

A. THE RATIONAL ACTOR MODEL

Thomas Schelling developed the rational actor model. The rational actor model is based on the cost of punishment associated with noncompliance multiplied by the likelihood of the punishment being implemented as the decision criteria.[21] This is weighted against the benefits of noncompliance. What this means is that a rational actor will evaluate the likelihood that he will be punished by an external actor and the capability of that external actor to punish for noncompliance against the benefits they will receive for noncompliance. If the likelihood that they will be punished is low or the capability of that actor to punish is low, then the external actor will have less influence on the rational actor's

[18] Schelling, 1966, 38–58.

[19] Wendt, 2005, 2–13.

[20] "Logic Model Development Guide," 2004, 1.

[21] Schelling, 1966, 38–58.

decision-making process. If there is more than one external actor, the rational actor will make the evaluation of the likelihood and capability to punish for each external actor. The rational actor model is about weighing costs and benefits of actions. This can be further illustrated by a simple equation.

(Eb*B) – (Ec*C) = Decision

Eb – expectation for receiving a benefit

B – the benefit itself

Ec – expectation of receiving a cost

C – the cost itself[22]

This equation or model will be used in a simple evaluation of a common situation faced by villagers in Afghanistan: being confronted by both the Taliban and International Security Assistance Forces (ISAF). In this example, both ISAF and the Taliban are looking to hold influence over the villagers. The villagers will weigh the benefits they will receive by supporting or not supporting ISAF or the Taliban. The weight of this analysis will be compared against the cost of noncompliance with either the Taliban or ISAF based on the Taliban and ISAF's likelihood and capability to punish or inflict costs. This model is being continuously applied to all decisions as the villagers "assess the values of costs and benefits every moment of every day."[23] This model is useful in gaining insight into the behavior of the villagers. This model is also useful for U.S. forces in explaining why some actions have been ineffective in Afghanistan.

This model can be used in explaining why there is "surprisingly weak evidence base for the effectiveness of aid in promoting stabilization and security objectives in Afghanistan."[24] The villagers who are receiving the benefits of the aid do not interpret the value to be greater than the costs the Taliban can inflict or

---

[22] Edward R. Evans and James R. Spies, *INSURGENCY IN THE HOOD: UNDERSTANDING INSURGENCIES THROUGH URBAN GANGS*, Thesis, Defense Analysis, Naval Postgraduate School (Monterey: Naval Postgraduate School, 2006).

[23] Evans and Spies, 2006, 17.

[24] Brooks, 2011.

the benefits the Taliban can give. The villagers are rational actors who are making rational decisions. In order for aid and development to support ISAF in Afghanistan, the benefit of the aid must outweigh the cost/benefit of Taliban influence or ISAF must be able to nullify the likelihood and capability of the Taliban to impart costs and benefits. If ISAF does not adequately address costs and benefits in the decision making process of the villagers, the villagers, as rational actors, will make decisions that may not support ISAF.

B. DR. GORDON MCCORMICK'S DIAMOND MODEL

Dr. Gordon McCormick's model, illustrated below in Figure 1, can be used to better understand counterinsurgency.[25]

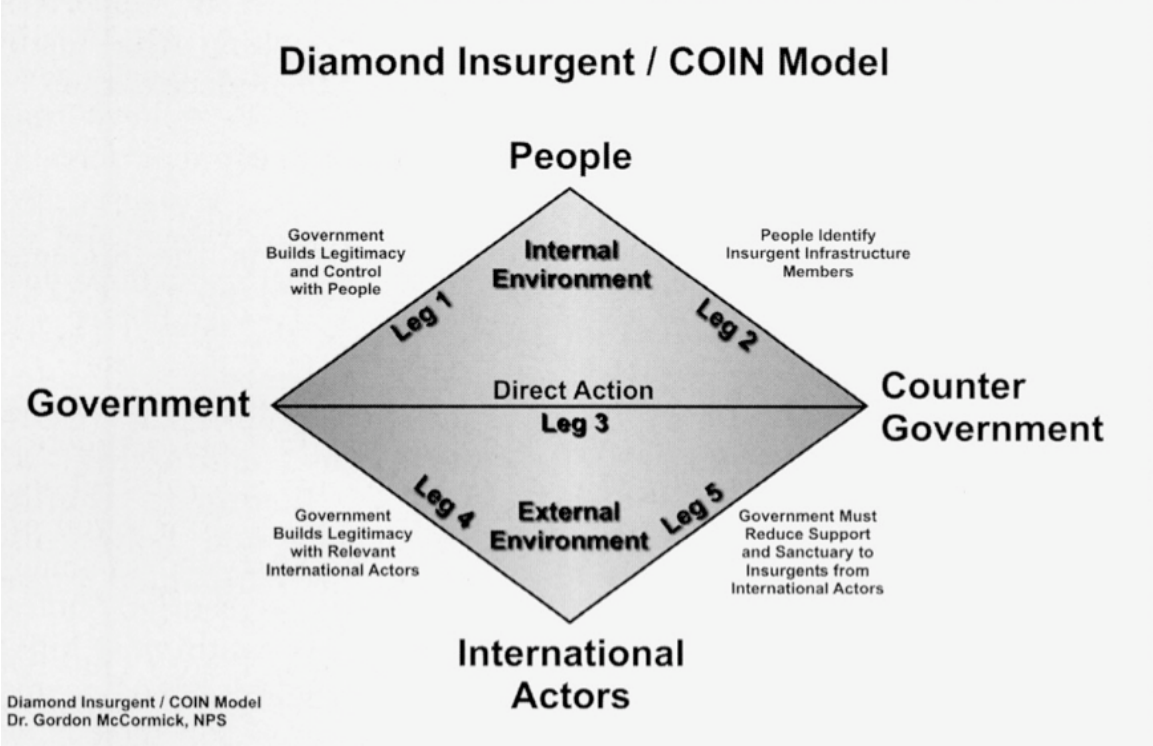


Figure 1. Dr. Gordon McCormick's Diamond Model[26]

[25] Wendt, 2005.

[26] Wendt, 2005.

Looking at the above model, at each point of the diamond there is an actor who plays a role in counterinsurgency: the government or state, the people or populace, the counter-government or insurgent, and the international actors. Five legs construct the relationships among these four actors. Leg One represents a course of action between the government and the people. Leg Two represents a course of action between the insurgency and the people. Leg Three represents a course of action between the government and the insurgent. Leg Four represents the relationship between the government and international actors. Leg Five represents a course of action between the insurgents and international actors.

The above diagram does not directly illustrate the feedback and support loop from the government and insurgency to the people and international actors.[27] The government holds a force and power advantage over the insurgency. The insurgency holds an information advantage over the government. So, the government needs the people to provide information and the insurgency needs the people to provide the resources (people, guns, and money) necessary to increase their power.[28] The support sought from international actors is the same for both the insurgency and the government; they both want logistics and legitimacy.

Based on this model, the government, operating on Leg One, needs to take action to increase its influence over the people to gain information that it can use to attack the insurgency (Leg Three), and more importantly the insurgency's relationship to the populace (Leg Two). The insurgency wants to accomplish the opposite. They need to attack the relationship between the government and the people (Leg One) by delegitimizing the government and strengthening their relationship with the populace through a mixture of coercion and persuasion.

---

[27] Mark C. Burke and Eric C. Self, *POPULATION ANALYSIS: A METHODOLOGY FOR UNDERSTANDING POPULATIONS IN COIN ENVIRONMENTS*, December 2008, http://www.dtic.mil/cgi-bin/GetTRDoc?Location=U2&doc=GetTRDoc.pdf&AD=ADA494013 (accessed March 15, 2011).

[28] Burke and Self, 2008.

The government wants to fortify ties to the international community (Leg Four) to maximize their incoming support and minimize the insurgents' ties to the international community. Again, the insurgency has similar objectives.

This is useful in developing metrics because the relationships between all actors are identified. The feedback mechanism is relatively transparent: is the government receiving information from the populace and is the insurgency receiving resources from the populace? These quantitative questions can be used to evaluate effects and outcomes. Is the populace giving more information to the government than they are giving resources to the insurgency? This feedback mechanism sparked initial interest in forming a dependent variable for the data analysis. The original plan was to use incoming tips, representing the government's receiving information, as the dependent variable. If more tips were received, this would correlate to increased stability. Unfortunately, it was quickly found that the data analysis was proving inconclusive. This will be covered in more detail in Chapter IV.

However, this is not to say the Diamond Model is not valid for determining metrics. The problems with the data analysis stemmed from incomplete information in the data. What were the circumstances under which the data was collected? What was the situation during which the data was collected? The validity of the data was the problem, not the Diamond Model. The Diamond Model can be used by units on the ground to help determine what effects they are having. The unit on the ground will know many of the circumstances around the increase or decrease in tips and will be aware of the situation to fully understand the context behind an increase or decrease in tips.

C. THE LOGIC MODEL

An important tool for developing useful MOEs in COIN is the logic model. According to the Kellogg Foundation, "A logic model is a systematic and visual way to present and share your understanding of the relationships among the resources you have to operate your program, the activities you plan, and the

changes or results you hope to achieve."[29] The logic model is composed of four elements: inputs, activities, outputs, and outcomes.

Inputs are defined as the resources available to "utilize towards doing the work."[30] Resources can be human, such as the number of troops available, the amount of money available, or equipment. Activities are what the organization does with the resources. "Activities are the processes, tools, events, technology, and actions…"[31] Outputs, by contrast, are the direct products of the activities. Outputs could be the number of high value targets (HVTs) killed or captured, the number of schools built, or the number of patrols conducted. Outcomes are the specific changes in the target's "behavior, knowledge, skills, status, and level of functioning."[32]

Below, in Figure 2, is an illustrative example of a logic model applied to common missions from Operation Iraqi Freedom.

Inputs	Activities	Outputs	Outcomes
30 soldiers 30 M4 Rifles 4 MRAPs	Cordon and Search of M145	1 – HVT Captured Cache discovered	Populace in M145 more supportive of GoI and USF
1 Civil Affairs Team 1 MRAP $500,000 in CERP	Contract a school to be built in M342	1 school with teachers and educated children	Populace more supportive of GoI and USF

Figure 2. Example Logic Model

29 "Logic Model Development Guide," 2004, 1.

30 "Logic Model Development Guide," 2004, 2.

31 "Logic Model Development Guide," 2004, 1.

32 "Logic Model Development Guide," 2004, 2.

Outcomes are the results of the operation. Per the example above, the outcomes are "more supportive to the Government of Iraq." A $500,000 school is not an outcome. A $500,000 school adequately staffed and producing educated children is not an outcome. A $500,000 school adequately staffed, producing educated children in an area, M145, *that is now more supportive to the Government of Iraq* (GOI) is an outcome. This represents a change to the target's, M145, behavior. Within the logic model, building a functioning school is an *output* not an outcome. A $500,000 school adequately staffed, producing educated children in an area that is still placing improvised explosive devices (IEDs) against U.S. Forces and the Government of Iraq is an outcome; the outcome is just negative to the Government of Iraq and U.S. forces. In this example and in nation assistance more broadly, measuring outputs is easier than measuring outcomes. Nevertheless, it is essential to focus on outcomes, because outcomes are results. Unfortunately, in many situations outputs are being measured instead of outcomes. The below passage by Amitai Etzioni is instructive and exemplifies tracking outputs, not outcomes.

> The newest way General Petraeus plans to measure success in the war in Afghanistan reminded me of what the government did when its campaign to persuade the public to stop smoking did not make much headway. It stopped counting how many people had had their last cigarette -- and started counting how many anti-smoking pamphlets it mailed.[33]

D. PERFECTION IS THE ENEMY OF GOOD ENOUGH: CORRELLATION VS. CAUSATION

1. Correlation

Correlation – "…a relationship exists between two factors—let's call them X and Y—but it tells you nothing about the direction of that relationship. It's possible that X causes Y; it's also possible that Y causes X; and it may be that X

---

33 Amitai Etzioni, *Beware of Generals Carrying Metrics*, September 30, 2010, http://www.huffingtonpost.com/amitai-etzioni/beware-of-generals-carryi_b_745343.html?ref=email_share (accessed June 13, 2011).

and Y are both being caused by some other factor, Z."[34] An easy example of correlation is between obesity and diet food. A trip to the supermarket may reveal that many obese people have shopping carts filled with diet food. However diet food is not the cause of obesity. Diet food may be correlated with obesity, but it is not the cause.[35]

2. Causation

Causation – A direct relationship exists between the independent variable and the dependent variable. There is a direction in the causal flow, "first x happens, and then y results."[36] An example of causation would be to look at the situation of the Allgemeine General Hospital in Vienna during the 1840s. The hospital was experiencing nearly a 10% death rate of mothers after delivery of their child. The cause of death was puerperal fever. Dr. Ignatz Semmelweis, through data analysis and one accidental occurrence, was able to trace the cause of the fever to the induction of cadaverous particles into the victim's bloodstream. The cadaverous particles were being carried on the hands of doctors who went directly from performing autopsies to delivering babies without washing their hands. The cure was as simple as using a chlorinated wash to disinfect the doctors' hands after performing an autopsy.[37]

3. Experience and Correlation

Malaria comes from the Italian words "mala aria," meaning bad air. The name came about as Roman Legions contracted malaria after traveling close to swamps. They believed the bad air coming from the swamps caused malaria.

---

[34] Levitt and Dubner, 2006, 10.

[35] The example of diet food and obese people was given in: Zachary Shore, *Blunder* (New York: Bloomsbury USA, 2008) 41.

[36] Shore, 2008, 41.

[37] Steven D. Levitt and Dubner J. Stephen, *Super Freakonomics* (New York: HarperCollins, 2009) 133–138.

Due to this belief, they began draining the swamps. The draining of swamps resulted in incidents of the disease subsiding.[38]

Now it is known that malaria is not spread by bad air but by a parasite carried by mosquitoes. Draining the swamps killed the mosquitoes, which in turn prevented the disease from infecting persons. The outcome of draining swamps was fewer cases of malaria.[39]

On one hand we can look at the Roman Legions and say they got lucky that cases of malaria decreased although they failed to identify the actual cause. On the other hand, we need to realize the Roman Legions were aware that areas without swamps did not result in cases of malaria. They saw correlation and acted on experience. The result of their actions was a success, fewer cases of malaria. In today's modern world, there are also fields where establishing causation is problematic, leaving action alone to correlation and experience.

4. Correlation Versus Causation in the Fields of Health and Medicine

T. Colin Campbell, PhD and Thomas M. Campbell II, MD further the discussion of correlation and causation in *The China Study*.[40] Campbell and Campbell's discussion of correlation versus causation is noteworthy, as causation in the medical and health field is nearly impossible to establish with absolute certainty.[41]

> The primary objective of research investigation (in health and medicine) is to determine only what is likely to be true. This is because research into health is inherently statistical. When you throw a ball in the air, will it come down? Yes, every time. That's physics. If you smoke four packs a day, will you get lung cancer? The answer is maybe. We know that your odds of getting lung

---

38 The example of the Roman Legions' response to malaria was in: Shore, 2008, 32.

39 Shore, 2008, 32.

40 T. Colin Campbell and Thomas M. Campbell II, *The China Study* (Dallas: BenBella Books, Inc, 2006, 38–40).

41 Campbell & Campbell, 2006, 38.

> cancer are much higher than if you didn't smoke, and we can tell you what those odds (statistics) are, but we can't know with certainty whether you as an individual will get lung cancer.[42]

Instead of absolute proof, researchers rely on probability, or statistical significance.[43] Statistical significance is used in analysis to determine if the observed effect is replicable under the same conditions or if the observed effect happened merely by chance.[44] Five percent or less probability is required for there to be statistical significance. This means that if the same conditions are present, the observed effect needs to be consistent ninety-five times out of one hundred for there to be statistical significance. 100 mice are injected with toxic X. Ninety-five of those mice develop liver cancer. The data from this experiment has statistical significance.

What needs to be stated is the ninety-five percent is an arbitrary, although widely accepted, number.[45] What also needs to be stated is the complexity of the conditions in which the effects are being observed. Campbell and Campbell were conducting a study on how diet relates to health. They list several problems in their study.

> Experimental limitations such as cost restraints, time constraints and measurement error are significant obstacles. Perhaps most importantly, food, lifestyle and health interact through such complex, multifaceted systems that establishing proof for any one factor and any one disease is nearly impossible, even if you had the perfect set of subjects, unlimited time, and unlimited financial resources.[46]

These problems bring perspective to trying to analyze relationships in nation assistance operations.

---

[42] Campbell & Campbell, 2006, 38.

[43] Campbell & Campbell, 2006, 40.

[44] Campbell & Campbell, 2006, 40.

[45] Campbell & Campbell, 2006, 40.

[46] Campbell & Campbell, 2006, 38.

5. Correlation Versus Causation and the Relationship of Dependent and Independent Variables

Alexander George and Andrew Bennett, in their book *Case Studies and Theory Development in the Social Sciences*, write "Highly general and abstract theories … which set aside intervening processes and focus on correlations between the 'start' and 'finish' of a phenomenon, are too general to make sharp theoretical predictions or to guide policy."[47] However, nation assistance missions have even more complexity, uncertainty, and much less control than studies from the field of health and medicine. Highly general and abstract theories may be the only theories, especially in nascent phases, to build on when trying to assess the relationship of the dependent variable to the independent variables.

The complexity of many situations in nation assistance, coupled with resource constraints, may preclude finding actual causation between dependent and independent variables. Process tracing to identify all the links in the causal chain may not be feasible or even necessary. Looking at an example from the health and medical field: despite years of study and millions of dollars, "Smoking has never been '100%' proven to cause lung cancer…"[48] It is unlikely that any one metric will ever be 100% proven to cause a specific outcome.

Members of the military are continuously reminded they need to be able to function under uncertainty. Finding causation between an independent variable and a dependent variable is unlikely. Finding correlation between an independent variable and a dependent variable is more probable. Pairing correlation to experience or research that shows the correlation to be a plausible relationship may be sufficient proof for action or decision.

---

[47] Alexander L. George and Andrew Bennett, *Case Studies and Theory Development in the Social Sciences* (Cambridge: MIT Press, 2005) 75.

[48] Campbell & Campbell, 2006, 41.

THIS PAGE INTENTIONALLY LEFT BLANK

III. WHEN IN DOUBT OF WHAT YOU SHOULD MEASURE, MEASURE WHAT YOU CAN

The process, if not the very idea, of measuring progress against an unconventional enemy is exceedingly complicated and is often fraught with potential pitfalls.[49]

In the process of data collection, the data had become an end unto itself.[50]

This chapter begins with historical problems in metrics, starting with body count in Vietnam. The chapter will continue to discuss the difficulties in developing metrics moving into the current conflicts in Iraq and Afghanistan.

A. INTRODUCTION

Colonel Gregory Daddis, from the history department at West Point, was tasked by Multi-National Corps – Iraq (MNC-I) in 2005 "to compose vignettes on past counterinsurgencies that might inform current U.S. military operations in Iraq."[51] Daddis and a colleague wrote two short essays. One was on the French-Indochina War and the other was on the Vietnam War. The papers were sent to Baghdad. After a few weeks, a Special Forces group commander, "responded to our (Daddis and his colleague) work by asking for any useful information on measuring progress in a counterinsurgency environment."[52] This group commander was in Iraq on his second tour. Daddis was surprised that this group commander who had just returned for another tour in Iraq "was struggling to delineate metrics…"[53]

Daddis' work on metrics and the Special Forces group commander's struggle with metrics came two years after former Secretary of Defense Donald

---

[49] Gregoray A. Daddis, *No Sure Victory: Measuring U.S. Army Effectiveness and Progress in the Vietnam War* (New York: Oxford University Press, 2011, x).

[50] Daddis, 2011, 10.

[51] Daddis, 2011, ix.

[52] Daddis, 2011, ix.

[53] Daddis, 2011, ix.

Rumsfeld wrote the following: “Today, we lack metrics to know if we are winning or losing the global war on terror.”[54] What has made metrics so difficult to determine? Looking back to World War II, it is hard to imagine Secretary of War, Henry Stimson, lamenting on the lack of metrics to determine whether or not the United States was winning or losing. Outcomes of operations in World War II could be seen on a map or judged from numbers on paper. In the sort of conventional fight of World War II, measuring progress was less ambiguous as capturing territory and killing the enemy both led to ultimate victory.[55]

> Terrain arguably served as the most visible scorecard. In fact, during the Normandy campaign, unit effectiveness and forward progress could be determined using a number of quantitative indicators – the number of troops or units ashore in France, the amount of territory under Allied control, the number of phase lines passed, or the number of Germans killed, wounded, or captured.[56]

Outside of conventional fights, such as the Vietnam War or our current involvement in Iraq and Afghanistan, metrics are not as clear. Capturing territory and killing the enemy may not lead to ultimate victory.[57] In 1965, as the war effort in Vietnam ramped up, the U.S. military began finding that “useful metrics for success or failure were not readily identifiable.”[58]

> With a ubiquitous enemy and no clearly defined front lines, U.S. soldiers and commanders struggled to devise substitutes for gauging progress and effectiveness. Their (U.S. military) conventional experiences offered few useful perspectives. Occupying terrain no longer indicated military success. The political context of fighting an insurgency complicated the process of counting destroyed enemy units or determining if hamlets and villages were secured or pacified. In short, the metrics for

---

[54] Rumsfeld, 2003.

[55] Daddis, 2011, 5.

[56] Daddis, 2011, 5.

[57] “In the end an insurgency is only defeated by good government which attracts voluntary popular support.” Robert Thompson, as cited in Daddis, 2011,12; In a counterinsurgency, the populace and the populace’s belief in and support of their government is the center of gravity, Sepp, 2005, 9.

[58] Daddis, 2011, 5.

assessing progress and effectiveness in World War II no longer sufficed for counterinsurgency operations.[59]

B. METRICS IN VIETNAM

1. Hearts and Minds and Body Count

When one thinks of Vietnam, the two common phrases that come to mind are "hearts and minds" and "body count." Hearts and minds came from General Sir Gerald Templer, Director of Operations and High Commissioner of Malaya, during the Malaya Emergency.[60] The term was spoken in the context "that success in COIN rests on the popular perception and this perception has an emotive ("hearts") component and a cognitive ("minds") component."[61] In 2006, the army released their counterinsurgency manual, FM 3-24. FM 3-24 goes farther with the definition developed by Templer.

> Once the unit settles into the AO, its next task is to build trusted networks. This is the true meaning of the phrase "hearts and minds," which comprises two separate components. "Hearts" means persuading people that their best interests are served by COIN success. "Minds" means convincing them that the force can protect them and that resisting it is pointless. Note that neither concerns whether people like Soldiers and Marines. Calculated self-interest, not emotion, is what counts. Over time, successful trusted networks grow like roots into the populace. They displace enemy networks, which forces enemies into the open, letting military forces seize the initiative and destroy the insurgents.[62]

"Hearts and minds" has some similarity of concept with another common phrase heard in counterinsurgency, "carrots and sticks." However, the phrase "hearts and minds" became more of a platitude that doling charity or being "nice" to the populace will somehow win their hearts and minds. In Vietnam, hearts

---

[59] Daddis, 2011, 5.

[60] Dave Dilegge, *Hearts and Minds*, October 21, 2007, http://smallwarsjournal.com/blog/2007/10/hearts-and-minds/ (accessed March 11, 2011).

[61] Dilegge, 2007.

[62] *Field Manual No. 3-24 COUNTERINSURGENCY* (Washington, DC: Headquarters Department of the Army, 2006).

and minds became the image of the "other war" – nation assistance operations.[63] Karen Guttierie further describes the other war as stability operations requiring a "whole community of effort, including coordination with governmental and non-governmental agencies from the host nation and the international community. This type of engagement characterized the economic assistance and other non-kinetic U.S. efforts in Vietnam…"[64]

The "other war" as name implies, always was a secondary effort to conventional operations.[65] The rejection of hearts and minds and the "other war" is seen with many decision makers in Vietnam eschewing programs of stability and counterinsurgency – "Grab 'em by the balls, and their hearts and minds will follow."[66]

The United States military put more effort into conventional operations, believing that their strength advantage would allow them to kill their way to victory.[67] It followed a strategy of attrition instead of looking to secure the populace – securing the populace, not killing the enemy, is the essential task in counterinsurgency.[68] But following the strategy of attrition in Vietnam, "Success

---

[63] Andrew F., JR. Krepinevich, *The Army and Vietnam* (John Hopkins University Press, 1986). 215–233.

[64] Karen Guttieri, "Metrics in Iraq's Complex Conflict Environment," in *The Three Circles of War*, 137–151 (Dulles, VA: Potomac Books, 2010, pp. 140–141).

[65] Krepinevich, 1986, 166.

[66] Stanley Karnow, *Vietnam: A History, Second Edition* (New York: Penguin Books, 1997) 450.

[67] Krepinevich, 1986, 166.

[68] Krepinevich, 1986, 164–168; COL(R) Kalev Sepp in an interview with PBS discussed a story about GEN Casey during his time as MNF-I Commander. The story is that a couple of months into his command he addressed his staff and said, "The number of insurgent deaths I'm receiving here is equal to or greater than the number two months ago you told me is the number of insurgents." Sepp goes on, "Gen. Casey started to get it right away: that this war was not going to be about victory through killing insurgents; that the Vietnam-style body count was not going to be the metric by which he could measure success in the country." *Frontline: Interview with Col. Kalev Sepp (Ret.)*, June 19, 2007, http://www.pbs.org/wgbh/pages/frontline/endgame/interviews/sepp.html (accessed November 03, 2011).

in counterinsurgency was made a function of the rate at which U.S. forces killed VC."[69]

Metrics are vital to assessing strategies in war.[70] What indicators are used will determine how operations proceed. "In Vietnam, the U.S. Marine Corps advocated, futilely, for the inclusion of crop production as a measure of success. (This indicator may have driven missions more towards supporting nation assistance operations.) The U.S. military focused on body counts rather than crop production, creating incentives to emphasize the kinetic fight rather than the 'other war.'"[71] Body count became the number one measure of effectiveness. Body count became the driving force behind operations.[72]

2. The Problems with Body Count

Great lengths were taken towards body count as the end goal. The United States' overwhelming firepower allowed for more tonnage of ordnance to be dropped on Vietnam than what was dropped in all of World War II. "By the time the Nixon administration signed a cease-fire agreement in January 1973, the United States had dropped on North Vietnam, an area the size of Texas, triple the bomb tonnage dropped on Europe, Asia and Africa during World War II."[73] In January 1969, during an operation in the Quang Ngai province, over 648,000 pounds of bombs and 2,000 rounds of artillery were used to kill 47 guerrillas.[74] During a three-month period in 1966, the ammunition expenditure equated to 1,000 rounds of artillery for one enemy killed.[75] Despite the inefficiency in the use of ammunition, attrition still prevailed as the way to victory.

---

[69] Krepinevich, 1986, 202.

[70] Guttieri, 2010, 137.

[71] Guttieri, 2010, 146.

[72] Krepinevich, 1986, 190–191.

[73] Karnow, 1997, 431.

[74] Krepinevich, 1986, 256.

[75] Krepinevich, 1986, 222.

"The body count quickly became the criterion for measuring success in Vietnam."[76] Securing the populace was thwarted in an effort to put more numbers (body count) on the board. However, body count did not factor in to the strategy of North Vietnam. "As General Giap noted, 'Every minute, hundreds of thousands of people die all over the world. The life or death of a hundred, a thousand, or tens of thousands of human beings, even if they are our own compatriots, represents really very little.'"[77]

Guttieri writes on measures of effectiveness in counterinsurgency, noting that falsely reading the environment is potentially disastrous.[78] The United States falsely read the environment. Incorrect reporting on body count furthered the false read of the environment.

The push to win through attrition led into problems with misreporting. The misreporting was both a result of willful misconduct, as in the case of the My Lai Massacre,[79] and negligence.

> The numbers never were accurate because some officers inflated their body counts to advance their careers or they simply guessed because guerrilla warfare in the jungles and rice paddies of Vietnam made counting bodies difficult. In 1967, members of the media began questioning whether numbers the military was providing were accurate, because the NVA/VC continually matched the U.S. escalation, fielding an army when their casualty numbers suggested they'd otherwise be unable to do so.[80]

FM 3-24 COUNTERINSURGENCY more plainly gives the reasons of body count as an ineffective metric.

---

[76] Krepinevich, 1986, 202.

[77] Krepinevich, 1986, 238.

[78] Guttieri, 2010.

[79] Colin Powell, *My American Journey* (New York, New York: Random House, Inc., 1995, 142–144).

[80] Ronald B., JR. Frankum and Stephen F. Maxner, *The Vietnam War Opposition in America Finds Meaning*, http://www.dummies.com/how-to/content/the-vietnam-war-opposition-in-america-finds-meanin.html (accessed March 03, 2011).

> They (body counts) actually measure very little and may provide misleading numbers. Using body counts to measure effectiveness accurately requires answers to the following questions:
>
> > How many insurgents were there at the start?
> >
> > How many insurgents have moved into the area?
> >
> > How many insurgents have transferred from supporter to combatant status?
> >
> > How many new fighters has the conflict created?
>
> Accurate information of this sort is usually not available.[81]

3. Problems with Other Metrics in Vietnam

Although hearts and minds and body count are the two most commonly thought of terms from the Vietnam War, the U.S. military did pursue many other types of metrics. Secretary Defense McNamara's advice to Military Assistance Command Vietnam (MACV) was "everything that was measurable should in fact be measured."[82] This resulted in massive amounts of data. Unfortunately, little was done to analyze the data and develop meaningful trends.[83] As a result, metrics from the ground were not nested with strategic objectives.

The problem with the developing metrics in Vietnam resulted from two main causes. "First, few officers possessed any real knowledge on how to gauge progress in an unconventional environment, particularly within the distinct setting of South Vietnam…(Second) They (officers) possessed even less understanding of the cultural landscape on which they were fighting."[84]

This resulted in nonsense reporting like in the following two examples.

---

[81] Field Manual No. 3-24 COUNTERINSURGENCY, 2006.

[82] Daddis, 2011, 10.

[83] Daddis, 2011, 10.

[84] Daddis, 2011, 9.

> Metrics are subject to distortion in the collection process. In his memoir of his days with a military advisory team in Vietnam, David Donovan reflected on the role of data in the Hamlet Evaluation System and Territorial Forces Evaluation System. These programs involved standardized lists of questions about troop strength, morale, numbers of houses with tin roofs, numbers of televisions, and the like. 'Meeting the deadline for submission of the report was the important thing,' he recalled, 'not accuracy.' The Vietnamese authorities and the U.S. district chiefs both understood that an A rating in the Hamlet Evaluation System put them in good light. Donovan wrote, "If I recall correctly, the month the infamous Tet offensive broke out, the country was reported to be over ninety-percent pacified.'[85]
>
> An example that best reflects the 'progress' being made involves several trips made to South Vietnam by Henry Kissinger, then a Harvard academic and adviser to New York governor Nelson Rockefeller. Upon visiting the province of Vinh Long in October 1965, Kissinger was told that 80 percent of the area had been pacified. When he returned to Vietnam the following July, Kissinger went again to Vinh Long and looked up the same official to check on how pacification was progressing. The man told Kissinger "enormous progress had been made" since his earlier visit: the province was now 70 percent pacified![86]

4. Conclusion

Inaccuracy in reporting, measuring the wrong metrics, and lack of effective analysis of what was reported all contributed to the ineffectiveness of metrics during the Vietnam War. Since there was a failure of metrics at the ground level, the metrics were not able to nest with strategic goals. The failure in developing effective metrics can be summed up in a conversation between COL(R) Harry Summers and a North Vietnamese colonel after the United States' withdrawal, "'You know, you never defeated us on the battlefield.' To which the Communist officer replied, 'That may be so, but it is also irrelevant.'"[87]

---

[85] Guttieri, 2010, 147.

[86] Krepinevich, 1986, 218.

[87] Karnow, 1997, 19.

C. METRICS IN IRAQ AND AFGHANISTAN

> ISAF must develop effective assessment architectures...to measure the effects of the strategy, assess progress toward key objectives, and make necessary adjustments. ISAF must identify and refine appropriate indicators to assess progress, clarifying the difference between operational measures of effectiveness critical to practitioners on the ground and strategic measures more appropriate to national capitals.[88]

General Stanley McChrystal said the above quote soon after taking command of the International Security Assistance Force (ISAF) in 2009. The war in Afghanistan had been going on for eight years before McChrystal took command. His quote means that process for developing metrics and products of metrics were not adequate. Therefore, eight years into the war, progress was being poorly assessed.[89]

1. Qualitative and Quantitative Metrics

Iraq and Afghanistan have created a need for strategists to create measures of effectiveness for NA missions. What has been created has had minimal success in determining what outcomes operations are having.[90] The Measures of effectiveness (MOEs) have been of two types, qualitative and quantitative. Military units have difficulty with the scholarly nature of qualitative assessments.[91] Quantitative assessments have a tendency to develop more indicators than military units can accurately report. Additionally, quantitative

---

[88] Jonathan Schroden, "Why Operations Assessments Fail: It's Not Just the Metrics," *Naval War College Review* 64, no. 4 (Autumn 2011): 89.

[89] Schroden, 2011, 90.

[90] Sean P. Sutherland, *Measuring Effectiveness in Conflict Environments*, September 2009, http://www.dtic.mil/cgi-bin/GetTRDoc?AD=ADA509874&Location=U2&doc=GetTRDoc.pdf (accessed January 27, 2011).

[91] David Wilson and Gareth E. Conway, "The Tactical Conflict Assessment Framerwork: A Short-lived Panacea," *Informaworld.com*, February 2009, http://www.informaworld.com/smpp/section?content=a909918221&fulltext=713240928 (accessed September 11, 2010).

assessments focus more on outputs (performance results) than on outcomes (effect results).

Within MOEs more broadly, there are challenges to using both qualitative and quantitative metrics. Some developers of qualitative assessments have argued that the intricacy and uniqueness of each environment in counterinsurgency, reconstruction, and stability operations do not allow for standardizing metrics. For example, a paper published by the Canadian Peacekeeping Press states that such a system "does not allow for the qualitative human phenomena upon which the military is coming in close contact…"[92] However, the average service member who interacts with the local populace needs a simple, standardized assessment system for efficiency.[93]

a. Tactical Conflict Assessment and Planning Framework (TCAPF)

The TCAPF is a qualitative assessment to "identify the causes of instability, develop activities to diminish or mitigate them, and evaluate the effectiveness of the activities in fostering stability in a tactical-level (brigade, battalion, or company) area of operations (AO)."[94] The TCAPF is a simple product that utilizes a four-step process consisting of collection, analysis, design, and evaluation.[95] The collection process is a two-part system. The first step is a simple questionnaire that is designed to be used by patrols to swiftly obtain a

---

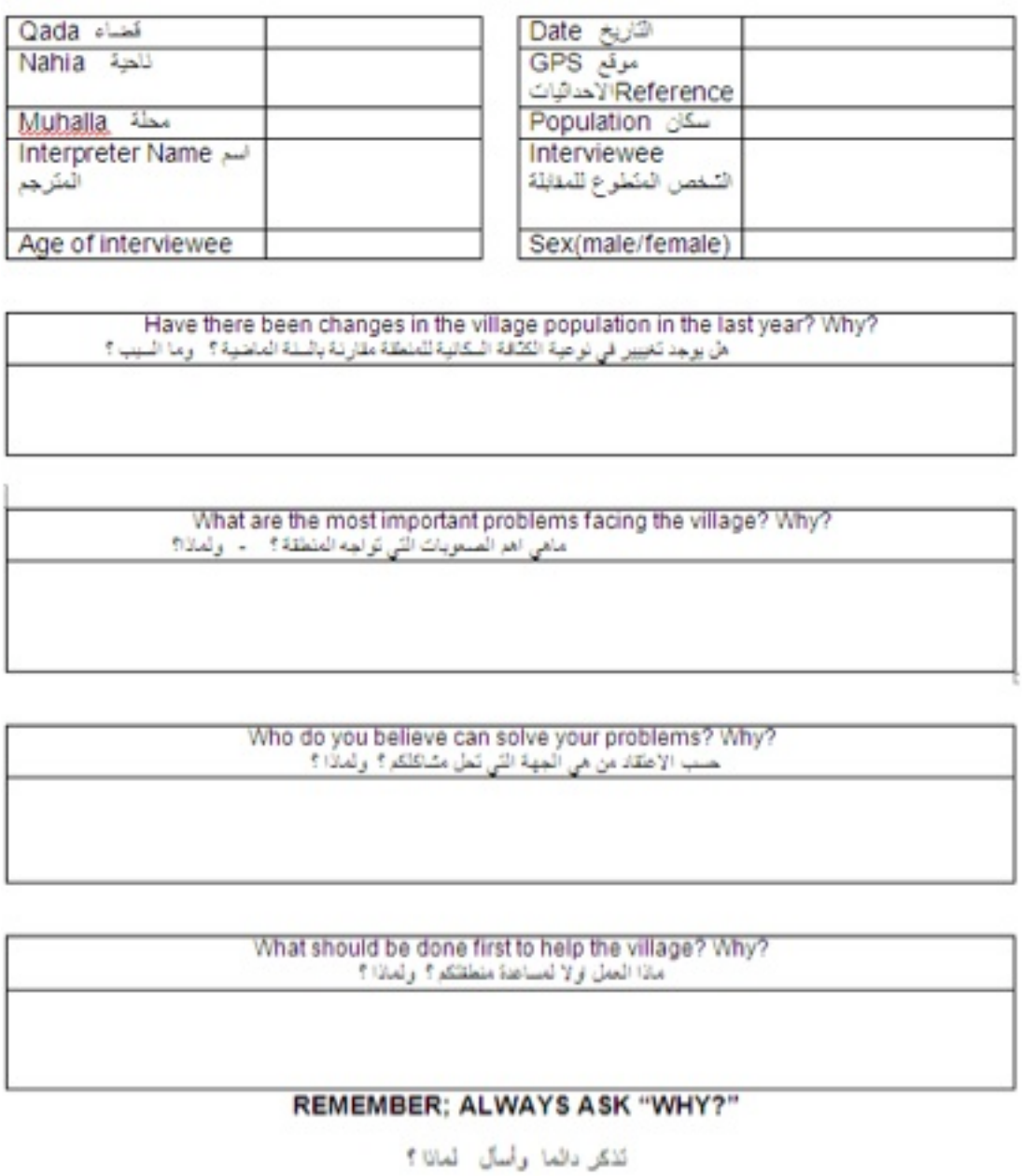
[92] Sarah Jane Meharg, "Measuring the Effectiveness of Reconstruction and Stabilization Activities," Pearson Papers, Measures of Effectiveness: Peace Operations and Beyond, Volume 10, Issue 1, ed. Ann Livingstone (Clementsport, Nova Scotia: The Pearson Peacekeeping Center, 2007).

[93] This is not to diminish the capabilities of members of the armed forces. This is the reality that the average service member on patrol is occupied with fulfilling the mission of the patrol and security concerns. This is the assessment of the authors of this thesis, who have read through numerous patrol reports on multiple deployments. The authors of this thesis assert that a simple, standardized assessment form produces the most consistent and accurate data.

[94] Michael Brooks, ed., *Measures of Effectiveness in Stability Operations*, ed. Michael Brooks, Vols. No. 10-41 (Ft. Leavenworth, Kansas: Center for Army Lessons Learned, 2010, 13).

[95] *Field Manual No. 3-07 STABILITY OPERATIONS* (Washington, DC: Headquarters Department of the Army, 2008, D10).

wide array of meaningful data in four questions. The four questions were selected to provide indicators of relative stability, causes of instability, level of support to the host nation government, and what can be done to aleviate causes of instability.[96] Below, in Figure 3, is an example of a TCAPF used in Iraq.



Qada قضاء		Date التاريخ	
Nahia ناحية		GPS موقع الاحداثيات Reference	
Muhalla محلة		Population سكان	
Interpreter Name اسم المترجم		Interviewee الشخص المتطوع للمقابلة	
Age of interviewee		Sex(male/female)	

Have there been changes in the village population in the last year? Why? هل يوجد تغيير في نوعية الكثافة السكانية للمنطقة مقارنة بالسنة الماضية ؟ وما السبب ؟

What are the most important problems facing the village? Why? ماهي اهم الصعوبات التي تواجه المنطقة ؟ - ولماذا؟

Who do you believe can solve your problems? Why? حسب الاعتقاد من هي الجهة التي تحل مشاكلكم ؟ ولماذا ؟

What should be done first to help the village? Why? ماذا العمل اولا لمساعدة منطقتكم ؟ ولماذا ؟

REMEMBER; ALWAYS ASK "WHY?"

تذكر دائما وأسأل لماذا ؟

Figure 3. Sample TCAPF from Iraq

[96] Field Manual No. 3-07 STABILITY OPERATIONS, 2008, D10.

The second part in the collection process is targeted interviews with key local stakeholders.[97] The second part is to act as a control to the first part of the collection process. Select individuals conduct the targeted interviews. This is in contrast to the first part that was designed for use at the lowest level. The targeted interviews go in to greater depth and "provide more detail on the causes of instability while helping determine how best to address those causes and measure progress toward that end."[98]

The information from the collection process is then analyzed. The information can be inputted as data into a geospatial software package or otherwise mapped to visually and graphically show the causes of instability, the sources of resiliency, and nodes for future targeting.[99]

Design and evaluation are the final two-steps in the four-step process. Design is the phase where programs are created to achieve effects on the identified causes of instability. Evaluation is the phase that tracks, compares, and measures the opinions of the populace.[100] The idea behind the TCAPF is the following:

> The TCAPF continually asks the people the same basic four questions so, over time, the changing opinion of the population can be tracked, compared, measured, and displayed. By not being subjective, the results of the questioning cannot be altered by those who ask questions or analyze results. Thus a longer-term narrative of the people is created that should ease situational awareness of new troops into theater and provide continuity of analysis and action.[101]

One of the authors of this paper, while serving as a BCT S9, used the TCAPF in Iraq. The results of the TCAPF varied greatly among the

---

[97] Field Manual No. 3-07 STABILITY OPERATIONS, 2008, D11.

[98] Field Manual No. 3-07 STABILITY OPERATIONS, 2008, D11.

[99] Field Manual No. 3-07 STABILITY OPERATIONS, 2008, D11.

[100] Field Manual No. 3-07 STABILITY OPERATIONS, 2008, D12.

[101] Brooks, 2010, 16.

battalions. The TCAPF is supposed to be a simple process. However, in many cases, the simple form of four basic questions was not filled out completely or correctly. Many forms had missing geographical data, rendering the information useless. The answers to the four questions were too vague to be useful. Even when the completed forms were properly filled out, when aggregated, the data was random and inconclusive.

Wilson and Conway note similar problems in their critique of the TCAPF, "the soldiers (using the TCAPF) are not as experienced as social scientists when it comes to a commitment to scientific rigor, nor are they experienced interviewers with a high level of ability to ask questions consistently and probe respondents in order to gain more relevant or more complete answers without biasing the data in any way."[102]

Note that the above critique does not go into analyzing the effectiveness of the four questions. The critique above is only on the application of the TCAPF in Iraq and Afghanistan.[103] This assumes that the four questions are capable of measuring outcomes.

b. Measuring Progress in Conflict Environments (MPICE)

Quantitative metrics are also difficult in counterinsurgency environments. Despite refuting the value of quantitative metrics in the Canadian Peacekeeping paper, the authors admit that "numbers count to the military."[104] However, quantitative assessments have their own challenges. Quantitative assessments have developed an intimidating number of indicators in efforts to measure all the intricacies in a counterinsurgency environment. For example, the Measuring Progress in Conflict Environments (MPICE) is a framework that contains over 800 individual metrics layered under multiple categories and

[102] Wilson and Conway, 2009.

[103] Wilson and Conway's article was based on experiences in Afghanistan.

[104] Meharg, 2007.

subcategories.[105] The unit has responsibility to choose which of these 800 metrics are relevant to their situation. Often, several hundred metrics are chosen. The results of trying to gather information on this many metrics can overwhelm soldiers on deployment, who have many other duties, and can result in misreporting, due to time constraints or lack of understanding of the metrics. This can lead to what Daddis cites in his book on metrics in Vietnam, "MACV – and much of DoD – went about measuring everything and, in a real sense, measured nothing."[106]

c. United States Agency for International Development's "F" Framework

Similarly, USAID's "F" Framework identifies almost 190 indicators for their strategic goal of promoting economic growth and prosperity.[107] In a report published by USAID, the author states, "existing systems are not providing adequate information to determine which programs have worked or how well they have worked in helping partner countries reach a sustainable path to rapid and broad-based growth."[108] The criticism of the "F" framework is that indicators measuring outcomes are supposed to serve as "bench marks of progress towards achieving higher-level outcomes."[109] This goal is not being achieved because the metrics are too many and too complicated.

---

[105] Michael Dziedzic, Barabara Sotirin and John Agoglia, Measuring Progress in Conflict Environments (MPICE): A Metrics Framework for Assessing Conflict Transformation and Stabilization, August 2008, http://www.usip.org/files/resources/MPICE%20Aug%2008.pdf (accessed November 21, 2011).

[106] Daddis, 2011, 10.

[107] USAID, *Measuring Effectiveness to Improve Effectiveness*, June 2010, www.countrycompass.com/_docs/policy_briefs/Briefing_Note_5_Effectiveness.pdf (accessed February 06, 2011).

[108] USAID, 2010.

[109] USAID, 2010.

d. RAND and Brookings Institute MOEs for Counterinsurgency

RAND and the Brookings Institute have also developed MOEs for counterinsurgency operations, particularly in Iraq and Afghanistan. The RAND report, *Measures of Effectiveness for the Information-Age Army*, has been criticized for being overly technical, utilizing calculus and math equations to measure data.[110] The Brookings Institute's Afghanistan Index earned the same critique, "far too comprehensive for field operators faced with spur of the moment decisions."[111]

e. Polling Data

Polling of local civilians has also become more widespread as field operators attempt to determine outcomes. However, polling data comes with its own set of problems. How people say they behave is known as declared preference, while how they actually behave is known as revealed preference.[112] Not surprisingly, the two preferences frequently differ. "There is often a vast gulf between how people say they behave and how they actually behave."[113] Because of this intrinsic shortcoming of human nature, polling data may not be so accurate.

f. David Kilcullen's MOEs from Counterinsurgency

Counterinsurgency expert David Kilcullen devoted a chapter in his latest book, *Counterinsurgency*, to a list of metrics useable by tactical-level units in Afghanistan. The metrics focus on outputs or, in Kilcullen's words, "detectable events in the environment that indicate progress toward, or away from, identified

[110] Sutherland, 2009.

[111] Sutherland, 2009.

[112] Levitt and Stephen, 2009, 7.

[113] Levitt and Stephen, 2009, 7.

goals."[114] Examples of what he terms "more useful metrics" are: voluntary reporting, transportation prices, and progress of NGO constructions projects. Kilcullen individually explains each one of the metrics and how it is an indicator of identified goals or outcomes. However, in the chapter's concluding paragraphs he states that the metrics explained in the previous pages "were also out of date the moment they were written down."[115] This last line points to the need for a system for developing metrics and not a prescriptive list.

2. Conclusion

The overarching problem with MOEs developed for Iraq and Afghanistan is that there is no system for *developing* metrics. Qualitative methods have created metrics too difficult for military units to operationalize. Quantitative methods require too many indicators for reporting and focus more on measuring outputs than outcomes. There is no system that allows units on the ground to develop their own specific metrics within a larger system that explains what needs to be measured and how to measure it.

[114] David Kilcullen, *Counterinsurgency* (New York, NY: Oxford University Press, 2010) 56.

[115] Killcullen, 2010, 76.

IV. DATA ANALYSIS – NUMBERS DO NOT LIE BUT THEY ARE OPEN TO INTERPRETATION

This chapter reviews our data analysis as well as our analysis of others' explanations of their data analysis in nation assistance operations. Prior to discussing our data and others' data, we discuss briefly the difficulty of selecting a viable dependent variable. Levels of violence and number of reported tips, commonly used dependent variables in data analysis, may not correlate to long-term stability. Long-term stability is a difficult concept and hard to capture as a variable. This will lead into the inconclusive results of past attempts at data analysis. Theses containing data analysis by David Beskow, Thomas O'Connell, Justin Gorkowski, and Jason Clark and Tracy Onufer will also be discussed.

A. COMMONLY USED DATA REPOSITORIES

1. Worldwide Incidents Tracking System (WITS) – Unclassified database. Compiles comprehensive data on "incidents in which subnational or clandestine groups or individuals deliberately or recklessly attacked civilians or noncombatants (including military personnel and assets outside war zones and war-like settings)."[116]

2. Tactical Integrated Ground Reporting (TIGR) – Classified database. Web-based geographic information sharing and reporting system, currently used by tactical-level units in Iraq and Afghanistan. TIGR receives and displays spatial combat reports entered directly by units. TIGR "merges data from established databases, notably Combined Information Data Network Exchange (CIDNE)."[117]

---

[116] *WITS Methodology and Criteria*, https://wits.nctc.gov/FederalDiscoverWITS/in (accessed October 11, 2011).

[117] David Beskow, *Spatial and Temporal Analysis of Tactical Integrated Ground Reporting*, Master's thesis (Monterey: Naval Postgraduate School, 2011).

3. Combined Information Data Network Exchange (CIDNE) – Classified database. "CIDNE was created to collect and analyze critical battlefield data to provide daily operational and intelligence community reporting relevant to a commander's daily decision-making processes."[118] CIDNE "has become the recognized source for analyzing enemy action, and is reliable for type, location, and frequency of significant enemy events."[119]

B. DIFFICULTY IN SELECTING A DEPENDENT VARIABLE IN NATION ASSISTANCE

1. Selection of the Dependent Variable

The first difficulty encountered in attempting data analysis for nation assistance missions is the selection of a viable dependent variable. In nation assistance missions, the goal of the operations is to promote long-term regional stability. Below is the definition of nation assistance from *Joint Publication 3-22, Foreign Internal Defense*:

> Civil or military assistance (other than foreign humanitarian assistance [FHA]) rendered to a nation by US forces within that nation's territory during peacetime, crises or emergencies, or war, based on agreements mutually concluded between the United States and that nation. NA operations support the host nation (HN) by promoting sustainable development and growth of responsive institutions. The goal is to promote long-term regional stability. NA programs include security assistance (SA), humanitarian and civic assistance (HCA), and foreign internal defense (FID). FID is the participation by civilian and military agencies of a government in any of the action programs taken by another government or other designated organization, to free and protect its society from subversion, lawlessness, insurgency, terrorism, and other threats to their security.[120]

---

[118] *CIDNE*, 2011, http://www.issinc.com/programs/cidne.html (accessed October 11, 2011).

[119] Beskow, 2011.

[120] Joint Publication 3-22: Foreign Interenal Defense (2010, I-1).

2. What is Representative of the Dependent Variable?

The dependent variable needs to represent the goal of promoting long-term regional stability. Herein lies the difficulty: what represents long-term regional stability? Some who have conducted data analysis for the conflicts in Iraq and Afghanistan have used violence against coalition forces, violence against civilians, or number of tips reported as the dependent variable.[121] Examples of the independent variables were the number of patrols, amount (both monetary and quantity) of aid or development projects, number of caches found or some sort of aggregated data sets like "friendly activity," "enemy activity," or "popular support." Unfortunately, due to the complexity of nation assistance and the complexity of the goal of nation assistance, choosing any of the aforementioned examples as a dependent variable may be lacking.

---

[121] Thomas J. O'Connell used the United Nations risk assessment level developed by the United Nations Department of Safety and Security (UNDSS) for the dependent varialbe at the district level. The independent variables were the amount and types of reconstruction and stabilization projects. Data was from the Afghanistan Country Stablility Picture (ACSP) and PRT project database. The dependent variable for the provincial level used the number of Taliban attacks against the populace. The Worldwide Incidicent Tracking System (WITS) was used as the database. The independent variables remained the same as in the district level analysis. Thomas J. O'Connell, *AFGHANISTAN RECONSTRUCTION - A QUANTITATIVE ANALYSIS OF THE INTERNATIONAL EFFORT*, March 2008, http://www.dtic.mil/cgi-bin/GetTRDoc?AD=ADA479990&Location=U2&doc=GetTRDoc.pdf (accessed October 10, 2011). Justin B. Gorkowski used an equation of CERP expenditure to normalized violence for the dependent variable. CERP data came from Iraq Reconstruction Management System (IRMS). SIGACT data came from U.S. Army Geospatial Center. The independent variables were labeled "needs" and "deserve." "Needs" was derived from village assessments. Deserve was derived from an equation composed of number of HVTs, number of tips reported, number of SIGACTs, and number and locations of media outlets. Justin B. Gorkowski, *A PENNY FOR YOUR THOUGHTS, A NICKEL FOR YOUR HEART: THE INFLUENCE OF THE COMMANDER'S EMERGENCY RESPONSE PROGRAM ON INSURGENCY*, December 2009, http://www.dtic.mil/dtic/tr/fulltext/u2/a514272.pdf (accessed October 10, 2011). Jason A. Clarke and Tracy L. Onufer conducted data analysis on the level of violence in Salinas, CA. The violence rate was used as the dependent variable. The independent variables were economy, population, housing, education, police force, prison influence, gang rivalry, social service programs, and community involvement. Jason A. Clarke and Tracy L. Onufer, *UNDERSTANDING ENVIRONMENTAL FACTORS THAT AFFECT VIOLENCE IN SALINAS, CALIFORNIA*, December 2009, http://www.dtic.mil/cgi-bin/GetTRDoc?Location=U2&doc=GetTRDoc.pdf&AD=ADA514256 (accessed October 10, 2011).

3. Historic Problems Representing the Dependent Variable in Data Analysis

The most commonly seen dependent variable has been represented by violence, either against coalition forces (SIGACTs pulled from CIDNE), the populace (pulled from WITS) or both. However, a decrease in violence may not mean progress towards long-term stability. "A decrease in enemy incidents might mean the government was in control but might also mean the insurgents were so established politically they no longer needed to fight."[122] David Kilcullen supports this idea. "Violence tends to be high in contested areas and low in government-controlled areas. But it is also low in enemy-controlled areas, so that a low level of violence indicates that *someone* is fully in control of a district but does not tell us who."[123]

Levels of violence also have potential to increase as coalition presence becomes greater. An increase of either troops in a given area, or the number of areas in which troops are operating, has potential to raise violence levels as there is a higher probability of enemy contact.[124] Kilcullen also notes the "observer effect" from increased troop presence. The observer effect results in an increase in reported violence as the increased troop presence also brings an increase in "eyes out on the ground observing and reporting violence."[125] An additional effect, Kilcullen notes in the *Accidental Guerrilla*, is that they fight us because we are there.[126] Kilcullen theorizes:

> Most of the adversaries Western powers have been fighting since 9/11 are in fact accidental guerrillas: people who fight us not because they hate the West and seek our overthrow but because we have invaded their space to deal with a small extremist element

[122] Daddis, 2011, 12.

[123] Kilcullen, 2010, 57–58.

[124] Kilcullen, 2010, 58.

[125] Kilcullen, 2010, 58.

[126] David Kilcullen, *The Accidental Guerrilla* (New York, NY: Oxford University Press, Inc., 2009) 263.

> that has manipulated and exploited local grievances to gain power in their societies. They fight us not because they seek our destruction but because they believe we seek theirs…[127]

Not only may additional troops increase violence through more probable contact with the enemy, the additional troops may increase violence by creating more "enemy." However, the main point of this section is that violence is not representative of long-term stability and therefore is not a good selection as a dependent variable.

5. A Viable Dependent Variable

A viable dependent variable for nation assistance will be one that represents long-term stability. Unfortunately, the complexity of long-term stability may preclude any one specific variable such as level of violence from being used. Long-term stability is composed of many factors and many factors are specific to individuals in the populace. "Political will, loyalty of the population, and an individual's sense of security cannot be accurately measured."[128] The coming chapters of this thesis will argue that ranking in the Failed State Index by the United States Institute of Peace may serve as the most viable dependent variable. In the coming paragraphs, we will continue to show the problems of data analysis.

C. DIFFICULTIES IN DATA ANALYSIS

1. Problems with Accuracy of Data Sets

For the conflicts in Iraq and Afghanistan, volumes of data have been collected. "Every day over 100,000 records of combat information are entered into various data bases in Afghanistan."[129] However, like all data, data from Iraq and Afghanistan "is only as accurate as the discipline, reporting standards, and

---

[127] Kilcullen, 2009, 263.

[128] Daddis, 2011, 14.

[129] Beskow, 2011.

priorities of those units reporting."[130] WITS has a disclaimer that states "The Worldwide Incident Tracking System (WITS) data is provided for statistical purposes only. The statistical information contained in WITS is based on factual reports from a variety of open sources that may be of varying credibility."[131] More importantly, WITS data is limited only to what the U.S. Government considers terrorist violence. Crime and other types of violence are not reported in WITS. Even if none of these problems existed, there are still problems.

When analyzing data, it is almost impossible to know the circumstances in which the data was collected. The data set may tell you exactly where, when, and how many improvised explosive devices (IEDs) were found, enemy direct attacks occurred, or tips called-in. However, the data does not tell you if more friendly or enemy troops moved into the area, if friendly or enemy troops changed their TTPs, or if there was a change in the civilian population. As a result, findings from the data can be scrutinized with questions that cannot be answered.

O'Connell found in his data analysis that there was an inverse relationship between local, small-scale security projects and Taliban attacks. "Local, small-scale security projects, rather than decreasing attacks, actually increased Taliban attacks, in some cases accounting for an amazing 76 percent of the increase."[132] What the increase does not show is why, or the other circumstances that may be responsible for the increase. Did the local, small-scale security projects result in the security forces patrolling in new areas or patrolling more aggressively? Did

---

[130] Beskow notes duplicate reporting, incomplete reports, and reports lacking geographical data as common problems. Beskow, 2011.

[131] WITS Methodology and Criteria, 2011.

[132] The 76 percent increase is misleading. The data analysis that O'Connell conducted resulted in a Multiple R value of .765. Data analysis more commonly uses R squared values, which are lower. Statistical significance is widely accepted at 95 percent. The Multiple R value of .765 really is not that high of a percent. Also, the Multiple R value is only explaining 76.5 percent of the variability in increased violence. This does not mean that security projects are increasing violence by 76 percent. O'Connell, 2008.

the security projects result in additional troops? Did more enemy move into the area during the time of the increase? Many antecedent conditions cannot be discerned in data analysis.

2. No Smoking Gun – The Results of Data Analysis

In this research, we will use the causes of instability (including, but not limited to, violence against civilians and direct and indirect attacks against coalition forces) as our dependent variable. Our goal is to find the variables that allow for either a decrease in this value or a leveling off of this variable at a stable value.

The analysis was focused on three areas/provinces in Afghanistan: Kunar and Kabul area, Kandahar province, and Helmand province. The data was obtained from WITS and declassified CIDNE data. The data was then normalized and aggregated into three categories, “Friendly Actions,” “Causes of Instability,” and “Indicators of Popular Support.” Friendly actions were all activities conducted by U.S. troops outside of a base. Causes of instability included but were not limited to violence against civilians and direct and indirect attacks against coalition forces. Indicators of popular support included but were not limited to called-in tips and caches found. The monetary value of Commander’s Emergency Response Program (CERP) projects were also included in the data analysis. Below, in Figures 4, 5, and 6, are the resulting graphs of these over time.

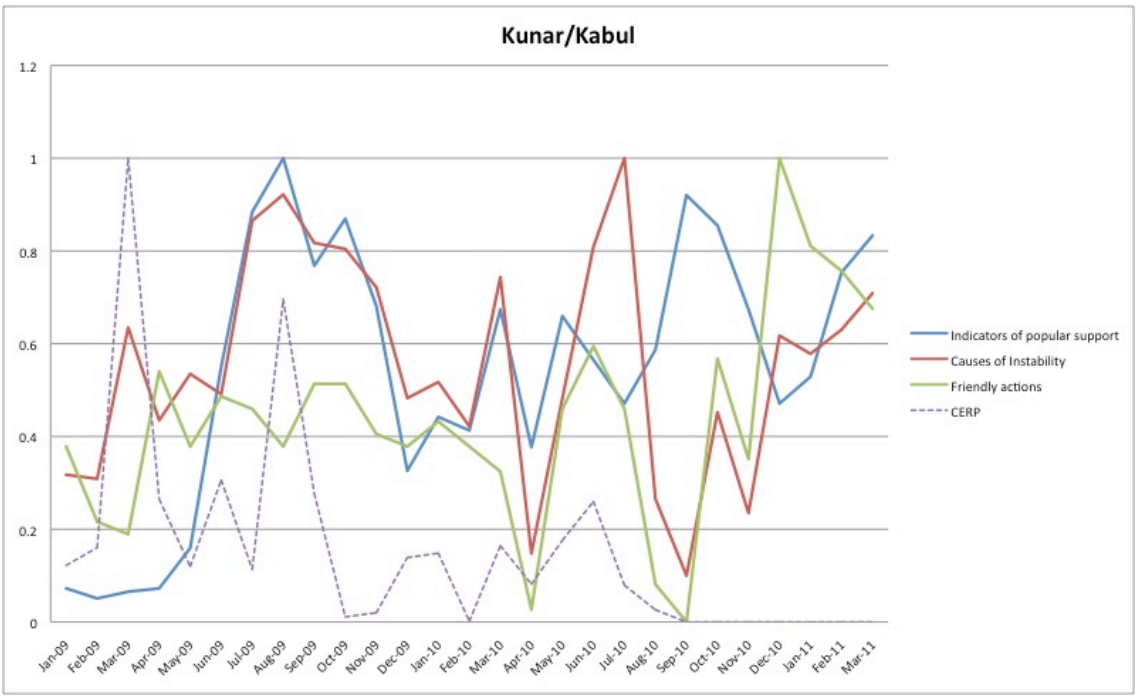


Figure 4. Kunar & Kabul Area Indicators of Popular Support, Causes of Instability, Friendly Actions and CERP Expenditure

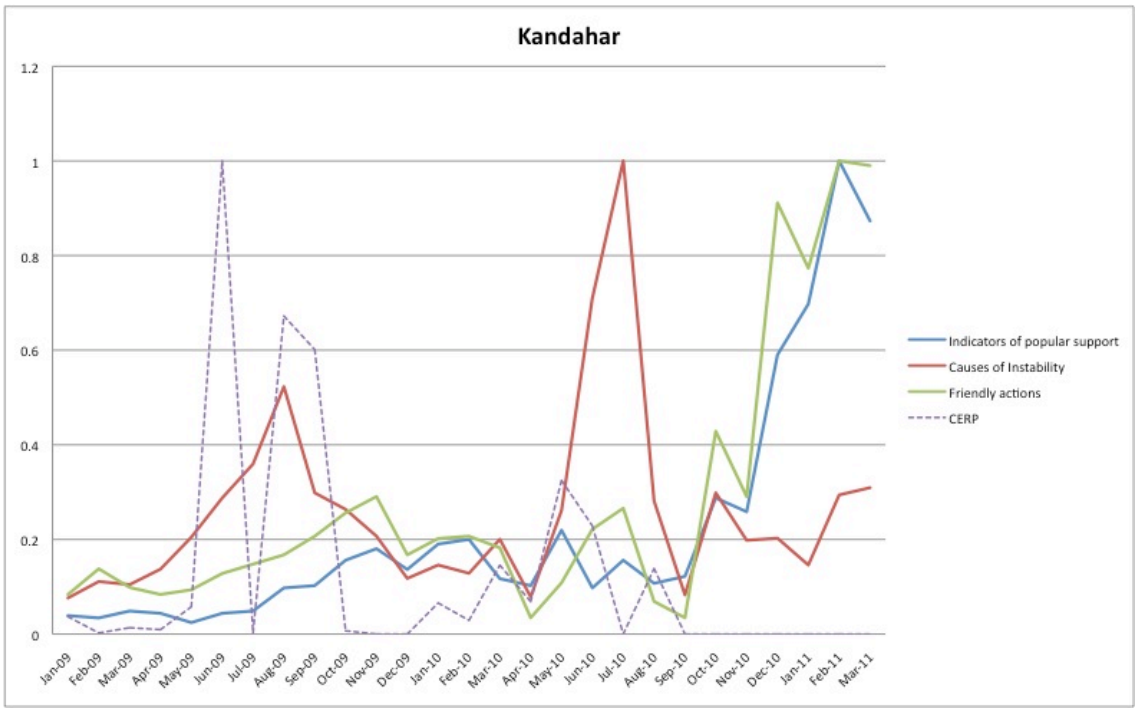


Figure 5. Kandahar Indicators of Popular Support, Causes of Instability, Friendly Actions and CERP Expenditure

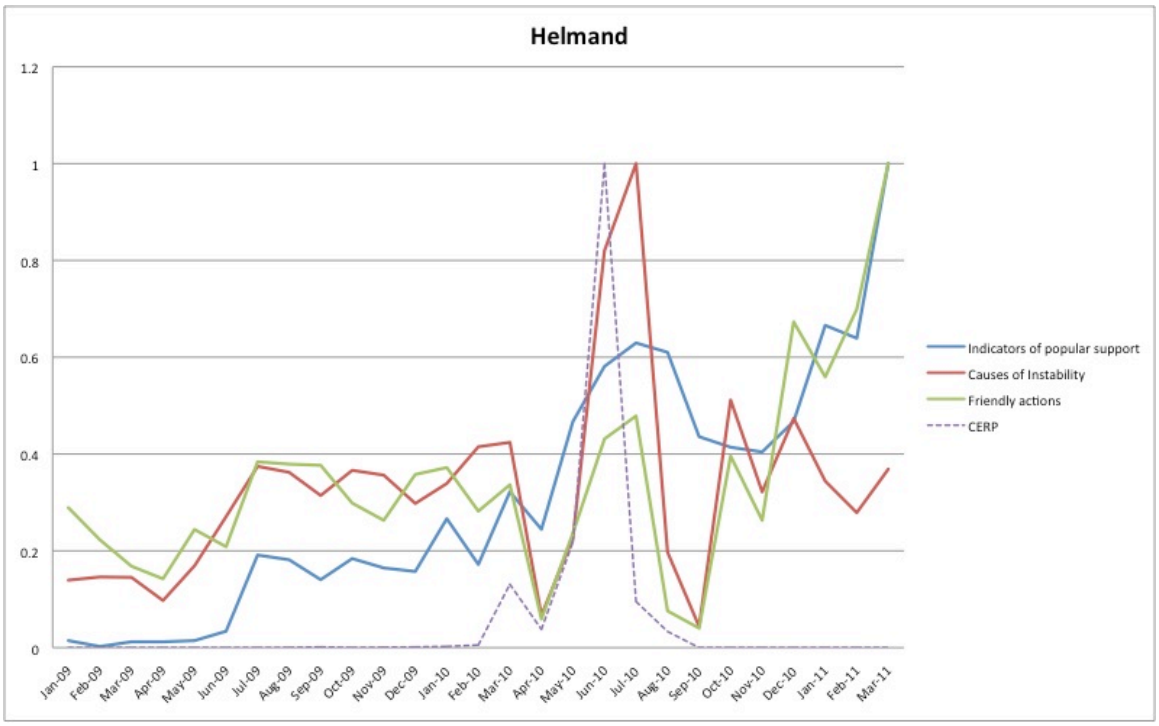


Figure 6. Helmand Indicators of Popular Support, Causes of Instability, Friendly Actions and CERP Expenditure

In the limited sampling above, CERP expenditure does not appear to have correlation to indicators of popular support. CERP expenditure does appear to have slight correlation to causes of instability, offset with a couple months' lag. From the data sets, an increase in CERP expenditure creates an increase in causes of instability a few months later. This seems counterintuitive. CERP should have been seen to increase indicators of popular support. Instead, the above graphs show the opposite may have occurred. Increased CERP spending may have no effect or negative impact on violence and causes of instability. Although counterintuitive, findings that show inverse relationships to CERP expenditure and stability are prevalent.[133]

[133] O'Connell (2008) found that expenditure on certain types of projects (education, medical care, and security) either had no effect on levels of violence or increased levels of violence. Gorkowski (2009) found that CERP expenditure in more "needing" but less "deserving" areas had weak correlation to increases in violence. Brooks (2011) reported on a conference of experts in Wilton Park in Britain who concluded that there is a 'surprisingly weak evidence base for the effectiveness of aid in promoting stabilization and security objectives' in Afghanistan."

Friendly actions appear to have a high degree of correlation to indicators of popular support in the Kandahar and Helmand data set above, with friendly actions closely mirroring indicators of popular support. This seems to suggest that just getting out and having presence amongst the populace creates popular support. However, this is a small sampling and the data set from the Kunar and Kabul areas appears significantly more random and less correlated between friendly actions and indicators of popular support.

Friendly actions and causes of instability appear to have a degree of correlation in the above three data sets. Friendly actions appear to lag behind causes of instability. This is somewhat intuitive; as instability increases, friendly actions will increase in an attempt to bring down the level of instability. Overall from this simple, basic data analysis, there appears to be nothing absolutely telling. The same is true in other's work with data analysis.

3. In a One-Chance, Zero Sum Game Where Every Option Has a Cost, What Probability Do You Need for Action?

Jason Clarke and Tracy Onufer, in their thesis, *Understanding Environmental Factors that Affect Violence in Salinas, California,* found that there is a 67 percent correlation between the unemployment rate and violence in Salinas.[134] "This means 67 percent of the time, if the unemployment rate increased or decreased, violence respectively increased or decreased."[135] Below, Figure 7, is the respective plot comparing 17 years of unemployment rates to rates of violence.

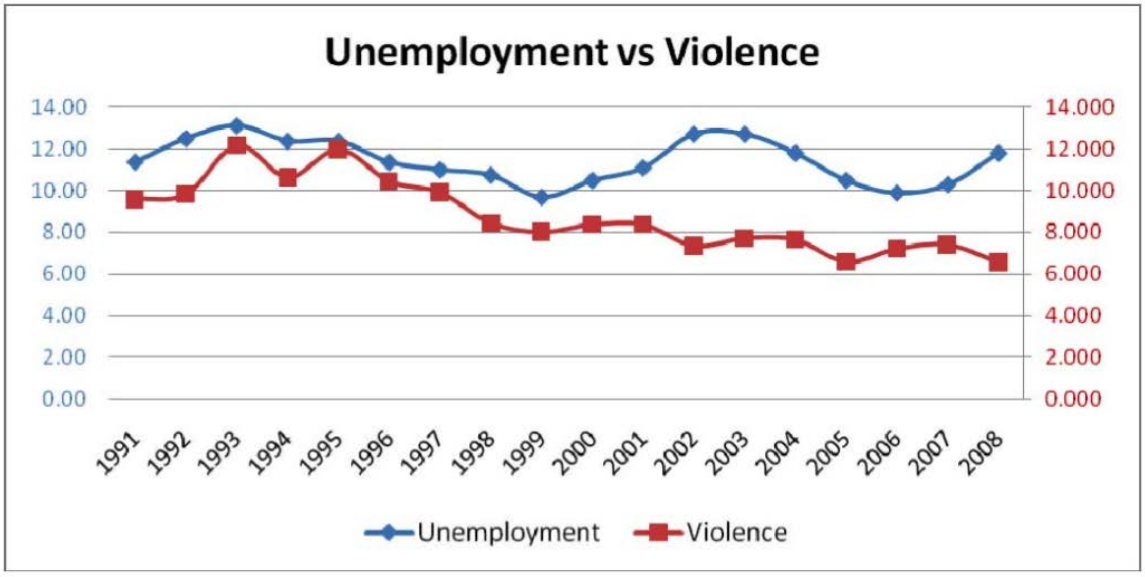


Figure 7. Unemployment Rate vs. Violence Rate in Salinas, CA[136]

First, Clarke and Onufer used the rate of violence as their dependent variable. Decreasing the rate of violence was their goal. Using rate of violence as a dependent variable works for this situation as the dependent variable

[134] Clarke & Onufer, 2009.

[135] Clarke & Onufer, 2009.

[136] Clarke & Onufer, 2009.

reflects the goal. To note from previous paragraphs, the exportability of the rate of violence as a dependent variable in nation assistance has difficulties.

Clarke and Onufer continue in their analysis and recommend that Salinas should develop a contingency plan for spikes in the unemployment rate.[137] The contingency plan "could include more police patrols as a proactive measure against violence, emergency funds to help the local citizens, food banks, and shelters, or job placement assistance to help laid-off workers."[138] However, every option has a cost and the city has a limited budget and limited number of personnel. Diverting money and personnel to one of Clarke and Onufer's recommended options takes money and personnel away from another program. This is a zero sum game. Is a 67 percent chance that violence will increase worth the cost of diverting resources? This situation becomes even more complex in nation assistance.

Justin Gorkowski in his thesis, *A Penny for Your Thoughts, A Nickel for Your Heart: The Influence of the Commander's Emergency Response Program on Insurgency*, analyzed CERP expenditure to violent acts in At Tameem, Iraq. In the analysis, Gorkowski found:

> Correlative analysis for violent acts and CERP money spent by village for all At Tameem villages reveals a value of (+) 0.302. This weak positive relationship shows that at the village level of analysis, the more CERP money that is spent, the more violent acts occur.[139]

[137] Clarke & Onufer, 2009.

[138] Clarke & Onufer, 2009.

[139] Justin B. Gorkowski, *A PENNY FOR YOUR THOUGHTS, A NICKEL FOR YOUR HEART: THE INFLUENCE OF THE COMMANDER'S EMERGENCY RESPONSE PROGRAM ON INSURGENCY*, December 2009, http://www.dtic.mil/dtic/tr/fulltext/u2/a514272.pdf (accessed October 10, 2011).

The below scatter-plot, Figure 8, shows the aforementioned relationship.

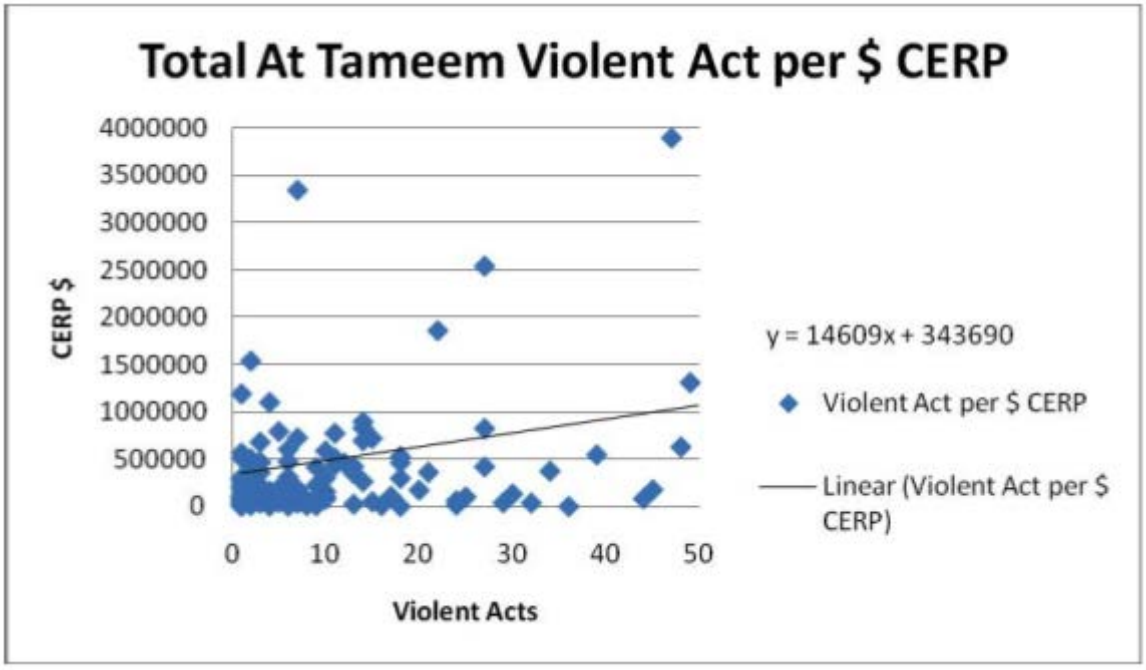


Figure 8. Scatter-plot of Violent Acts per CERP $ Spent in At Tameem

Gorkowski suggests that CERP expenditure is actually increasing violence. However, the correlation between CERP expenditure and increasing violence is only 30.2 percent. This means that only 30.2 percent of the variability in increasing violence is explained by CERP expenditure. Remembering that the widely accepted percent for statistical significance is 95 percent, this is weak correlation and correlation does not mean causation.

Does CERP money still get spent in hostile areas? These are tough questions and black-and-white answers are hard to find. What makes matters even more difficult is that the goal of long-term stability is not immediate. "The peace and security timetable is measured in years or decades. Development progress, if it comes at all, is measured in generations."[140] It is tough to find answers in data analysis.

[140] Brooks, 2011.

THIS PAGE INTENTIONALLY LEFT BLANK

V. OPERATIONALIZING THE FAILED STATE INDEX FOR MEASURING OUTCOMES AT THE TACTICAL LEVEL

A. INTRODUCTION

One of the goals of the United States in Iraq and Afghanistan is to protect the populace. In any counterinsurgency, a subset of nation assistance, the populace's support for their government is the center of gravity of the overall conflict.[141] Ultimate success is achieved by protecting the populace.[142]

The method being used towards this goal resembles nation building. Utilizing this assumption, we can then look to metrics used to measure outcomes in nation building. The Failed State Index developed by the Fund for Peace has developed metrics for this purpose. However, the metrics they are using are for measuring outcomes at the strategic level. This chapter will operationalize the Failed State Index for use by a tactical unit.

First, this chapter will start with an explanation of nation building as this activity is directly related to addressing the Failed State Index indicators. Second, this chapter will outline the Failed State Index developed by The Fund for Peace. The twelve indicators in the index will be fully explained to show how the indicators are measured. Third, the chapter will show how a tactical unit's operations in Sadr City addressed many of the indicators in the Failed State Index. Fourth, the indicators from the Failed State Index will then be operationalized into the Tactical Outcome Assessment and used in the case study of Sadr City to measure the tactical unit's outcomes.

---

[141] Sepp, 2005, 9.

[142] General Petraeus reiterated this while addressing NATO in July 2010 shortly after replacing General McChrystal as the new head of U.S. and NATO forces in Afghanistan, "You must do everything possible to protect the population." *US chief Petraeus vows to protect Afghan civilians*, July 1, 2010, http://www.bbc.co.uk/news/10472555 (accessed March 11, 2011). This idea is also stated in: Elliot Cohen, Conrad Crane, Jan Horvath and John Nagle, "Principles, Imperatives, and Paradoxes of Counterinsurgency," *Military Review*, March-April 2006: 52.

B. NATION BUILDING

The RAND Corporation, in a report titled *The Beginner's Guide to Nation Building* published in 2007, defines nation building as an operation that "involves the use of armed force as part of a broader effort to promote political and economic reforms, with the objective of transforming a society emerging from conflict into one at peace with itself and neighbors."[143] The most successful examples of nation building are Germany and Japan following World War II. The nation building efforts in Japan and Germany were "aimed to engineer major social, political, and economic reconstruction."[144] These efforts set a standard of success that has not been replicated.[145]

From the above paragraph, we can see that nation building is more than providing stability. Nation building is engineering major social, political, and economic reconstruction of a state. As a result, the essential elements of a nation building mission are composed of establishing and increasing the capacity of the targeted nation's military, police, rule of law, and governance, and improving and affecting humanitarian relief, economic stabilization, democratization and development.[146] These essential elements are also what the Failed State Index's twelve indicators are measuring.

C. THE FAILED STATE INDEX

1. Introduction to the Failed State Index

The Failed State Index was developed by The Fund for Peace, "an independent, nonpartisan, 501(c)(3) non-profit research and educational

---

[143] James Dobbins, Seth G. Jones, Crane Keith and Cole Degrasse Beth, *The Beginner's Guide to Nation Building*, 2007, http://www.rand.org/pubs/monographs/MG557.html (accessed August 31, 2011).

[144] James Dobbins, et al., *America's Role in Nation-Building: From Germany to Iraq*, 2003, http://www.rand.org/pubs/monograph_reports/MR1753.html (accessed August 31, 2011).

[145] Dobbins, et al., 2003.

[146] Dobbins, et al., 2007.

organization that works to prevent violent conflict and promote sustainable security."[147] The Fund addresses policy-level as well as ground-level issues emanating from weak and failing states. They have worked in over fifty countries with governments, international organizations, academics, journalists, civil society networks, and the private sector.[148] Every year, they rank 177 countries across twelve indicators. The Failed State Index is published annually by Foreign Policy magazine.

The Failed State Index only concerns itself with sovereign states recognized by membership in the United Nations. For this reason, Taiwan, the occupied territories of Palestine (these territories are included in the assessment of Israel), and Kosovo are not included. There are also additional states that are not included due to insufficient data.[149]

[147] *About the Fund for Peace*, 2011, http://www.fundforpeace.org/global/?q=aboutus (accessed August 21, 2011).

[148] The Fund for Peace, 2011.

[149] The Fund for Peace, 2011.

2. The Twelve Indicators of the Failed State Index

Figure 9 lists the 12 indicators in the Failed State Index.

Social Indicators
1. Mounting Demographic Pressures
2. Massive Movement of Refugees or Internally Displaced Persons
3. Legacy of Vengeance-Seeking Group Grievance or Group Paranoia
4. Chronic and Sustained Human Flight
Economic Indicators
5. Uneven Economic Development along Group Lines
6. Sharp and/or Severe Economic Decline
Political and Military Indicators
7. Criminalization and/or Delegitimization of the State
8. Progressive Deterioration of Public Services
9. Suspension or Arbitrary Application of the Rule of Law and Widespread Human Rights Abuse
10. Security Apparatus Operates as a "State Within a State"
11. Rise of Factionalized Elites
12. Intervention of Other States or External Political Actors

Figure 9. Failed State Indicators[150]

These twelve indicators are further broken down, on average, into fourteen sub-indicators or measures for each of the twelve indicators. Mounting demographic pressures contain sub-indicators or measures like disease control, food supply, population density, and infant mortality. The sub-indicators can be put into one-line questions that can be answered with either quantitative or qualitative data.[151]

---

[150] The Fund for Peace, 2011.

[151] The Fund for Peace, 2011.

3. Sources of Data for the Failed State Index

The data for the Failed State Index comes from news articles, essays, magazine articles, speeches, and government and non-government reports. Blogs and other forms of social media are not included in the data. Thousands of reports are compiled daily; millions of documents are downloaded annually from a search landscape of 115,000 online English language publications.[152]

The media is then subjected to a scan using proprietary software. Boolean phrases on indicators from the Conflict Assessment System Tool (CAST) framework retrieve relevant material while filters built into the software eliminate irrelevant and erroneous documents. Next, quantitative data from the United Nations High Commission of Refugees (UNHCR), World Health Organization (WHO), United Nations Development Program (UNDP), Transparency International, World Factbook, Freedom House, World Bank, and other reputable, reliable sources are brought into the Fund's proprietary software package. Then, a qualitative review of each indicator for each country is conducted to compare the results of the media search and quantitative data input. These three sources (media search, quantitative data, and qualitative review) are used in the methodology to triangulate the data to produce the final scores for the Failed State Index.[153]

4. Scoring the Failed State Index

The scores for the Failed State Index come from scaling (from 0-10) the aggregated, then normalized data for each of the twelve indicators. An algorithm is used for the scoring. A score of zero is the most stable. A score of 10 is the least stable. The 177 countries can receive a score anywhere between 0 and 120, with a score of 120 being the most unstable. Somalia, Chad, and Sudan (the three lowest-ranking states) have scores of 113.4, 110.3, and 108.7,

[152] The Fund for Peace, 2011.

[153] All information in this paragraph came from: The Fund for Peace, 2011.

respectively. The three highest-ranking states, Finland, Norway, and Sweden, have respective scores of 19.7, 20.4, and 22.8. The Fund for Peace ensures the validity of the scores. "This multi-stage process has several layers of scrutiny to ensure the highest standards of methodological rigor, the broadest possible information base including both quantitative and qualitative expertise, and the greatest accuracy."[154]

5. Uses for the Failed State Index

The Failed State Index provides validated information in an easy to access format to many different agencies.

- Government uses - early warning and to design economic assistance strategies that can reduce the potential for conflict and promote development in fragile states
- Military uses - strengthen situational awareness, enhance readiness, and apply strategic metrics to evaluate success in peace and stability operations
- Private sector uses - calculate political risk for investment opportunities
- Multinational organizations and other entities use - modeling and gaming, management of complex organizations, and for conflict-risk assessment
- Educators use - train students in analyzing war and peace issues by blending the techniques of information technology with social science
- Rated country uses - self-assessment to gauge their own stability and performance on objective criteria[155]

Overall, the use of the Failed State Index is to identify and diagnose problems. Identification and diagnosis is the first step in strengthening weak and failing states. This can be used as an early warning by "more reliable policymakers"[156] to implement assistance strategies that "can act to prevent

[154] The Fund for Peace, 2011.

[155] The Fund for Peace, 2011.

[156] "The more reliable policymakers" was a term from: The Fund for Peace, 2011.

violent breakdowns, protect civilians caught in the crossfire, and promote recovery."[157] Policies must focus on the five core institutions: military, police, civil service, system of justice, and leadership. The assessment methodology used in the Index should continuously monitor what effects a policy is having. Monitoring will allow the policymakers to make informed decisions regarding strategic choices for weak or failing states. The goal is to prevent failed states and assist in the recovery of failed states.

6. Characteristics of Failed States

"States can fail at varying rates through explosion, implosion, erosion, or invasion over different time periods."[158] A failed state will have one or more of the following characteristics:

- Loss of physical control of its territory or a monopoly on the legitimate use of force
- Erosion of legitimate authority to make collective decisions
- Inability to provide reasonable public services
- Inability to interact with other states as a full member of the international community

Risk Elements from the twelve indicators:

- Extensive corruption and criminal behavior
- Inability to collect taxes or otherwise draw on citizen support
- Large-scale involuntary dislocation of the population
- Sharp economic decline
- Group-based inequality
- Institutionalized persecution or discrimination
- Severe demographic pressures
- Brain drain
- Environmental decay

---

[157] The Fund for Peace, 2011.

[158] The Fund for Peace, 2011.

7. Conclusion

Operationalizing the Failed State Index is the first step in preventing failed states. The Failed State Index identifies problem states. Once a state is identified as failing, it is then up to the international community to decide whether or not their interests in the failing state and the results of the failing state are worth the cost of action.

The action can be nation building and the United States will most likely be involved to some extent. The RAND Corporation states, "Nation-building, it appears, is the inescapable responsibility of the world's only superpower."[159] Nation building, in many instances, takes place in conflict environments that require armed forces. The Army is the branch of service most directly involved with nation building as they deploy units on nation assistance missions. The following case study demonstrates how an Army unit at the tactical level conducted nation assistance operations while deployed to OIF 08-09.

D. SADR CITY CASE STUDY

1. Background and History of Sadr City

Hassan Shama, Sadr City District Council Chair, once asked, "Do you know that in some parts of the city, we have more than 30 family members living in an area of 1000 square feet?"[160] Sadr City is a slum; a Shi'a slum that suffered under a repressive Sunni government and then fought against a Sunni insurgency.

Sadr City is one of the nine katis or districts that make up the Amanat of Baghdad.[161] Inside the city of Baghdad, Sadr City lies to the northeast, approximately 5 miles east of the Tigris River. On a map, it looks roughly like a

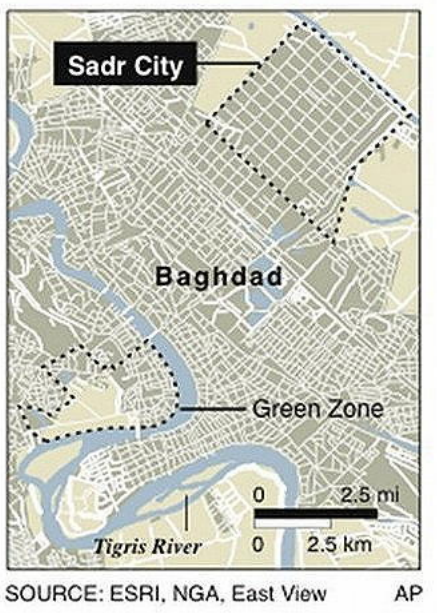
---

[159] Dobbins, et al, 2007.

[160] Anita McNaught, *Winning the Peace in Iraq by Rebuilding Sadr City*, June 26, 2008, http://www.foxnews.com/story/0,2933,371764,00.html (accessed May 27, 2011).

[161] Amanat refers to the term, "inside the city." *Republic of Iraq District Government Field Manual Version 2*, PDF (USAID, 2007).

baseball diamond with the legs of the square measuring approximately three miles. In these nine square miles, population estimates have placed the number of inhabitants anywhere between a low of 2 million and a high of 3.2 million.[162] Baghdad has a population of 5.75 million.[163] Using the low population estimate, Sadr City has a population density higher than Calcutta or Hong Kong.[164] Figure 10 is a map of Sadr City.



[162] The high of 3.2 million came from an interview of MAJ Humphreys of 3rd BCT, 4th Infantry Division in an article by Michael Totten. Michael Totten, *Sadr City After the Fall*, April 8, 2009, http://www.michaeltotten.com/archives/2009/04/sadr-city-after.php (accessed May 25, 2011). The low of two million came from an article by Glen Tschirgi. Glen Tschirgi, *Sadr City's Lesson for Gaza*, January 13, 2009, http://www.americanthinker.com/2009/01/does_sadr_city_have_a_lesson_f.html (accessed May 25, 2011). However, despite sources stating the population as over 2 million, it is hard to believe. Hong Kong and Calcutta are stacked with high-rise residential buildings. Sadr City's residences average less than four stories. But if there are 30 people living in 1000 sqft dwellings, the reported population may be correct.

[163] *Middle East: Iraq*, May 17, 2011, https://www.cia.gov/library/publications/the-world-factbook/geos/iz.html (accessed May 22, 2011).

[164] Damian Cave, *Sadr City starts to turn around, posing new challenges - Africa & Middle East - International Herald Tribune*, February 08, 2007, http://www.nytimes.com/2007/02/08/world/africa/08iht-sadr.4525914.html (accessed May 27, 2011).

Figure 10. Map of Sadr City[165]

Iraqi Prime Minister Abdul Karim Kassem built Sadr City in the late 1950s, as public housing for the city's poor.[166] The streets were lined in a grid pattern and 125,000 residences were built. The individual residences were approximately 1,500 square feet.[167] Since the 1950s, the original 1,500 square foot residences have been subdivided and subdivided again as the population grew. The subdivision and growing population have only hastened the deterioration of the infrastructure, which had already faced neglect for three decades under the government of Saddam Hussein and the resulting instability following the U.S. invasion in 2003.[168] Mike Davis, in *Planet of Slums*, writes of places like Sadr City as volcanoes waiting to erupt.[169]

Colonel Peter R. Mansoor and Major Mark S. Ulrich write on counterinsurgency and list three prerequisites for an insurgency to exist. They break the prerequisites down into three categories:

- Vulnerable Population – a vulnerable population is one with social, political, economic, or security related grievances
- Leadership for Direction – There must be a person, group, or idea to mobilize the vulnerable populace
- Lack of government control – The government must be non-responsive or overly repressive. Therefore, the government does not have legitimacy[170]

---

[165] Michael Totten, *Sadr City After the Fall*, April 8, 2009, http://www.michaeltotten.com/archives/2009/04/sadr-city-after.php (accessed May 25, 2011).

[166] *Sadr City*, May 22, 2011, http://topics.nytimes.com/top/news/international/countriesandterritories/iraq/sadr_city/index.html (accessed May 22, 2011).

[167] Cave, 2007.

[168] Cave, 2007.

[169] As cited in Peter W. Singer, *The Future of National Security, By the Numbers*, May 22, 2011, http://www.brookings.edu/articles/2011/05_national_security_singer.aspx (accessed May 22, 2011).

[170] Colonel Peter R. Mansoor and Major Mark S. Ulrich, "Linking Doctrine to Action: A New COIN Center-of-Gravity Analysis," [PDF] *Military Review* (September-October 2007).

Sadr City met all three prerequisites. The population was neglected under Saddam Hussein. The sectarian violence and lack of security after the fall of Saddam furthered the population's vulnerability. Muqtada Al Sadr emerged to provide leadership and direction. The unstable conditions following the fall of Saddam created a void of government control. The conditions in Sadr City were ripe for an insurgency.

2. Nation Building at the Tactical Level

This was the situation that confronted one Army unit when they assumed responsibility of Sadr City in late 2007. The first priority was to establish security. The plan to do this was similar to what was used in Gaza. In 2008, U.S. forces began constructing a concrete wall that ran the length of Sadr City, cutting off the Southern quarter of city. The wall, known as the Gold Wall, as it ran along Route Gold, limited the Mahdi Army's freedom to maneuver and protected the populace from their influence. "Once the citizens behind the Gold Wall were confident of continuing protection from Jay'sh al-Mahdi (JAM) reprisals, businesses re-opened, security improved dramatically and actionable intelligence from the population soared."[171] Also during this time, Muqtada al Sadr, called for a ceasefire, and stood down JAM.[172]

With security in place, further development and reconstruction could occur. U.S. forces and other American agencies began working on a range of projects from health to education. Twenty-two roads were nominated for repaving and improvement. Sixteen sewer mains in the city were cleaned out to eliminate the festering pools of waste that once polluted the town and its market areas.[173] In this year, 2008, $44 million dollars in U.S. aid was spent inside Sadr

---

[171] Tschirgi, 2009.

[172] Michael R. Gordon, *War Over Wall Persists in Sadr City Despite Truce*, May 15, 2008, http://www.nytimes.com/2008/05/15/world/middleeast/15wall.html (accessed May 31, 2011).

[173] Cave, 2007.

City.[174] Markets that once were vacant under sectarian violence were now bustling with business.[175] Parks were being filled with children. Normalcy was returning to Sadr City.[176]

Assistance was not limited to security, economic, and essential service assistance. Advisors from the U.S. State Department mentored local government officials. The mentorship was in the nascent phase and having government officials just attend a meeting was progress. As progress continued, mentorship of the local government officials expanded with the goal to assist them in performing the basic governmental functions of communicating between the different levels of government, making decisions, and executing a budget.[177]

E. EVALUATING THE TACTICAL UNIT ON THE FAILED STATE INDEX

1. Operationalizing the Failed State Index

The Failed State Index was designed for use at the strategic level. The index is looking at stability of the state. The evaluation of the indicators will need to be modified in order to function at the tactical level.

The modification relies on filtering the indicators from the Index through the logical lines of operations (LLO) in Army doctrine *FM 3-24 COUNTERINSURGENCY.*[178] The result of this process is the Tactical Outcome Assessment. The Tactical Outcome Assessment, shown in Figure 11, is the operationalized Failed State Index for use at the tactical level.

---

[174] *Reconstruction: Brig. Gen. Talley, Mr. Nazar Al-Sultan, Mr. Bass, Jan. 15*, January 15, 2009, http://www.usf-iraq.com/?option=com_content&task=view&id=24945&Itemid=131 (accessed May 31, 2011).

[175] Totten, 2009.

[176] Totten, 2009.

[177] The information in this paragraph is from eyewitness accounts of the author who served in Sadr City from 2009-2010 and personally knew the State Department Officials involved in the mentoring.

[178] Field Manual No. 3-24 COUNTERINSURGENCY, 2006, 5-3.

The Failed State Index	LLOs	Tactical Outcome Assessment
Mounting Demographic Pressures	Combat Operations/ Civil Security Operations	Inc/Dec Security
Massive Movement of Refugees or Internally Displaced Persons		
Legacy of Vengeance-Seeking Group Grievance or Group Paranoia	Host Nation Security Forces	Inc/Dec Governance
Chronic and Sustained Human Flight		
Uneven Economic Development along Group Lines	Essential Services	Inc/Dec Economic Opportunities
Sharp and/or Severe Economic Decline		
Criminalization and/or Delegitimization of the State		Inc/Dec Access to Essential Services
Progressive Deterioration of Public Services	Governance	
Suspension or Arbitrary Application of the Rule of Law and Widespread Human Rights Abuse		Inc/Dec Rule of Law
Security Apparatus Operates as a "State Within a State"	Economic Development	Inc/Dec Movement/Number of Refugees or Internally Displaced Persons
Rise of Factionalized Elites		
Intervention of other States or External Political Actors		Inc/Dec Return/Flight of Influential Individuals

Figure 11. Tactical Outcome Assessment

The most important part of these indicators is determining what metric will represent each individual indicator. The tactical unit doing the measuring must make this determination. Ideally, battalion or special operations task force (SOTF) staffs will determine the representative metrics. However, it may be more practical for brigade or joint special operations task force (JSOTF) staffs to handle this effort as their staffs are more robust than battalions and SOTFs.

Some metrics representative of the indicators are obvious—refugees and internally displaced persons (IDPs) are represented by refugees and IDPs. Other indicators need more deduction. What represents economic opportunities in a specific area? More stores opening in Jamilla market may represent increased economic opportunities in Sadr City, Iraq. More goats per family may represent increased economic opportunities in Helmand, Afghanistan. A decrease in the tangle of electric wires running from private generators or illegal taps that tangle the city streets and are ripped down by the antennas of military

vehicles on patrol may mean an increase in access to essential services in Baghdad, Iraq. The need for private generators and illegal wiretaps diminish, as the Iraqi government is able to improve its ability to provide electricity to its citizens. This is reflected by the decrease in "spider webs" of electrical cords shown in Figure 12.

Figure 12. Electrical Wires in Adhamiyah[179]

Metrics representative for most of the indicators can be found by simply asking, "What does right look like?" To determine an indicator for rule of law, the correct question is, "What should a functioning judicial system look like in this area?" The answer may be a tribal elder handling disputes without corruption or accepting bribes. What does increased security look like? In some areas, it may look like functioning checkpoints. What should increased access to essential services look like? In Baghdad, Iraq one of the sub-indicators would look like

[179] Totten, 2009.

city-provided power, void of illegal taps. The answers to these questions will identify indicators and provide a basis for measurement. If the answers cannot be seen or found, they are not good indicators.

The questions should be simple. Does this village have a tribal elder that is handling disputes? If yes, how effective is the tribal elder on a scale of one to ten? Does this town have functioning checkpoints? If yes, how effective are the checkpoints on a scale of one to ten? Keeping the questions to yes and no responses with a variable scale standardizes the process and allows for aggregating the scores like in the Failed State Index. To note, these scores should not be the only information on the area; qualitative assessments should accompany.

Once metrics representative of the indicators have been found and measured, numerical values will be assigned. The ten-point scoring system, the same method used by the Failed State Index, will be used. Each indicator will be assigned a value from one to ten. The scores for each indicator will then be aggregated to determine a single number for their area. Liberty is being exercised in aggregating the numbers. The numbers being aggregated are not cardinal numbers. However, coming up with a single number as a score is beneficial as it aids in showing trends—whether the situation is improving or worsening.

The Tactical Outcome Assessment, although not tested, should be more consistent in determining outcomes for the following reasons:

- This system is not prescriptive and, therefore, should not lead to problems with metrics discussed in Chapter III.
- This system credits the unit on the ground with being the expert in its area and mandates unit involvement in determining what will be measured.
- This system develops area specific metrics to accurately represent the ten indicators, instead of dictating to the unit on the ground specific metrics that may not apply.

The use of this system should not overly task military units. A typical BCT or JSOTF staff—through mission analysis, intelligence preparation of the battlefield (IPB), and establishing commander's critical information requirements (CCIR) —already has the knowledge base to develop metrics representative of the seven indicators for their OE. Next, developing a collection plan for the metrics is not anything different than what BCT staffs are already doing. Subordinate units who will execute the collection plan are not being tasked with anything outside of what they are already routinely tasked to do. This system bears no additional cost.

The Tactical Outcome Assessment is something that can be included into the unit's priority intelligence requirements (PIR). The PIR reporting standard will remain the same except for the additional requirements. This system, for the above reasons, should be effective at determining outcomes that are accurate (within a unit's ability to collect), informative, and determine where and how to spend time and money.

2. Evaluating a Tactical Unit in Sadr City with the Operationalized Failed State Index

This section serves as an example of how the Failed State Index, operationalized to the tactical level, can evaluate outcomes in a tactical unit's OE. Information has been gained from the above case study. Not all information is available to complete the assessment of outcomes. However, this example will demonstrate the basics of the system.

Sadr City is a densely populated urban area. What does right look like in a densely populated urban area? Should there be a functioning sewage system? Should there be a functioning police or security force? Questions like these need to be asked to determine metrics to represent the seven indicators.

Sub-indicators for 1. Increasing/Decreasing Security for this case study will be:

- Degree change in the number of SIGACTs from previous reporting period (0-10), 0 equals great reduction in SIGACTs, 5 equals no change, 10 equals great increase
- Degree change in the number of tips resulting in actionable intelligence from previous reporting period (0-10), 0 equals great increase in tips, 5 equals no change, 10 equals great decrease
- Degree change in the number of people in markets from previous reporting period (0-10), 0 equals great increase, 5 equals no change, 10 equals great decrease
- Degree change in the number of people in parks from previous reporting period (0-10), 0 equals great increase, 5 equals no change, 10 equals great decrease

Sub-indicators for 2. Increasing/Decreasing Governance for this case study will be:

- Degree change in the effectiveness of scheduled meetings from previous reporting period (0-10), 0 equals great increase in functionality, 5 equals no change, 10 equals great decrease

Sub-indicators for 3. Increasing/Decreasing Economic Opportunities for this case study will be:

- Degree change in market place activity from previous reporting period (0-10), 0 equals great increase in functionality, 5 equals no change, 10 equals great decrease
- Degree change in the number of new businesses from previous reporting period (0-10), 0 equals great increase in functionality, 5 equals no change, 10 equals great decrease

Sub-indicators for 4. Increasing/Decreasing Access to Essential Services for this case study will be:

- Rate of progress of construction projects from previous reporting period (0-10), 0 equals great increase in functionality, 5 equals no change, 10 equals great decrease
- Degree change in sewage on the streets from previous reporting period (0-10), 0 equals great reduction in SIGACTs, 5 equals no change, 10 equals great increase

Utilizing the above sub-indicators and 10-point scoring system, the case study will provide the following result illustrated below in Figure 13:

1. I/D Security (0 best, 10 worst)	**2**

1.1 Degree change in the number of SIGACTs from previous reporting period (0-10), 0 equals great reduction in SIGACTs, 5 equals no change, 10 equals great increase	2
1.2 Degree change in the number of tips resulting in actionable intelligence from previous reporting period (0-10), 0 equals great increase in tips, 5 equals no change, 10 equals great decrease	3
1.3 Degree change in the number of people in markets from previous reporting period (0-10), 0 equals great increase, 5 equals no change, 10 equals great decrease	1
1.4 Degree change in the number of people in parks from previous reporting period (0-10), 0 equals great increase, 5 equals no change, 10 equals great decrease	1
2. I/D Governance (0 best, 10 worst)	**4**
2.1 Degree change in the effectiveness of scheduled meetings from previous reporting period (0-10), 0 equals great increase in functionality, 5 equals no change, 10 equals great decrease	4
3. I/D Economic Opportunities (0 best, 10 worst)	**2**
3.1 Degree change in market place activity from previous reporting period (0-10), 0 equals great increase in functionality, 5 equals no change, 10 equals great decrease	1
3.2 Degree change in the number of new businesses from previous reporting period (0-10), 0 equals great increase in functionality, 5 equals no change, 10 equals great decrease	3
4. I/D Access to ESS (0 best, 10 worst)	**3**
4.1 Rate of progress of construction projects from previous reporting period (0-10), 0 equals great increase in functionality, 5 equals no change, 10 equals great decrease	3
4.2 Degree change in sewage on the streets from previous reporting period (0-10), 0 equals great reduction in SIGACTs, 5 equals no change, 10 equals great increase	2
Tactical Outcome Assessment (1. + 2. + 3. + 4.)	**11**

Figure 13. Tactical Outcome Assessment Filled Out

The Tactical Outcome Assessment does rely on subjectivity of the one conducting the assessment. There is a possibility that those responsible for

completing the assessment will be tempted to report more favorably. However, it is the unit that has created the metrics it believes are capable of measuring outcomes in their OE. This should create a vested interest to mitigate against inaccurate reporting.

3. Tactical Outcomes Nesting With the Broader Strategy in U.S. Nation Assistance Missions

The above example of measuring outcomes at the tactical level will readily nest with the broader strategy in U.S. nation assistance missions. The outcomes above are measuring stability in accordance with the Failed State Index. The goal of U.S. nation assistance missions is stability.

The example from the above case study in 2008 shows increasing stability for the area. Utilizing the Failed State Index, which measures stability at strategic levels: Iraq in 2007 was listed as the second most unstable country in the world; Iraq in 2008 was listed as the fifth most unstable country in the world; Iraq in 2009 was listed as the sixth most unstable country in the world; Iraq in 2010 was listed as the seventh most unstable country in the world; Iraq in 2011 was listed as the ninth most unstable country in the world. For the last 5 years Iraq has become more and more stable. The above case study nests with these results, although significantly more case studies would be needed to show that this method of measuring outcomes does nest with the broader strategy of U.S. nation assistance missions.

THIS PAGE INTENTIONALLY LEFT BLANK

VI. CONCLUSION AND RECOMMENDATIONS

A. CONCLUSION

This thesis has investigated two vital problem areas: first, what models are helpful in developing metrics that measure outcomes in nation assistance operations and, second, what models are ineffective in determining metrics that measure outcomes in nation assistance operations? It also addressed the two initial research questions: *What metrics should tactical level military units employ for measuring outcomes in nation assistance (NA) operations? How should a unit's measures of effectiveness in its specific operating environment connect to metrics of its command and ultimately nest with the broader strategy in U.S. NA operations?* The Tactical Outcome Assessment developed in Chapter V is a significant step toward final answers of the latter two questions.

1. Theories and Models Useful in Determining Metrics That Measure Outcomes

This thesis has found models and definitions that are useful in determining metrics that measure outcomes in nation assistance operations. The rational actor model, used by Thomas Schelling, is helpful as it explains the motivation of the populace.[180] Dr. McCormick's Diamond Model is useful in analyzing an area as it contains a feedback mechanism that can be used to explain increasing or decreasing success as the population support moves from the insurgent to the state or from the state to the insurgent.[181] The logic model is necessary for understanding the differences between inputs, throughputs, outputs, and outcomes.[182] These terms are necessary in developing metrics. Correlation and causation are two additional terms that need to be understood for developing metrics.

---

[180] Schelling,1966, 38–58.

[181] Wendt, 2005: 2–13.

[182] "Logic Model Development Guide", 2004, 1.

a. Rational Actor Model

Thomas Schelling's rational actor model is based on the cost of punishment associated with noncompliance multiplied by the likelihood of the punishment being implemented as the decision criteria.[183] This is weighted against the benefits of noncompliance. What this means is that a rational actor will evaluate the likelihood that he will be punished by an external actor and the capability of that external actor to punish for noncompliance against the benefits they will receive for noncompliance. If the likelihood that they will be punished is low or the capability of that actor to punish is low, then the external actor will have less influence on the rational actor's decision-making process. If there is more than one external actor, the rational actor will make the evaluation of the likelihood and capability to punish for each external actor. The rational actor model is about weighing costs and benefits of actions.

What this means for the development of metrics is that it explains the motivation of the populace. The populace will act a certain way based on how they interpret the costs and benefits of action in a specific situation. If we see more women out by themselves, wearing more western style dress in an area of traditional Taliban control, we can infer that the degree of Taliban control has lessened, as those women no longer believe there is a cost that the Taliban can impose.

b. Dr. Gordon McCormick's Diamond Model

Dr. McCormick's Model simplifies counterinsurgency to the interactions between the four actors, the government, insurgency, populace, and international actors.[184] The model lays these actors out in a diamond pattern with one of the actors at each point in the diamond. The lines that connect the points of the diamond pattern represent the relationships amongst the actors. The model contains a feedback and support loop to measure the populace's

[183] Schelling, 1966, 38–58.

[184] Wendt, 2005.

support for the government or insurgency: support for the government is seen by the populace supplying the government information on the insurgency; support for the insurgency is seen through the populace supplying the insurgency with people, guns and money.

For developing metrics, this model gives two hard metrics. The first is the number of actionable intelligence items received from the populace. The second is the amount of aid and support given by the populace to the insurgency. The first metric is easier to track through the standard reporting procedures units in Iraq and Afghanistan are currently using. However, the second metric, the amount of aid and support given by the populace to the insurgency, is harder to identify.

c. The Logic Model

The logic model is composed of four elements: inputs, activities or throughputs, outputs, and outcomes. Inputs are defined as the resources available to "utilize towards doing the work."[185] Resources can be human, such as the number of troops available, the amount of money available, or equipment. Activities are what the organization does with the resources. "Activities are the processes, tools, events, technology, and actions…"[186] Outputs, by contrast, are the direct products of the activities. Outputs could be the number of high value targets (HVTs) killed or captured, the number of schools built, or the number of patrols conducted. Outcomes are the specific changes in the target's "behavior, knowledge, skills, status, and level of functioning."[187]

The terms used in the logic model are needed for developing metrics. Without knowledge of the terms, it is easy to develop metrics that instead of developing outcomes end up measuring outputs, throughputs, or inputs. The latter three do not matter in determining what effect operations are

185 "Logic Model Development Guide," 2004, 2.

186 "Logic Model Development Guide," 2004, 1.

187 "Logic Model Development Guide," 2004, 2.

actually having. They only measure the efficiency of unit performance and not the effectiveness of unit performance.

d. Correlation vs. Causation

Causation is a term used to refer to a direct relationship between a dependent variable and an independent variable. Correlation is a term used to refer to a relationship that is not direct but exists to some degree between a dependent variable and an independent variable. In correlation, the degree to which the relationship exists factors in to whether the effect seen on the dependent variable is replicable or happened by chance.

Correlation is based on the percent of variability in the relationship determined through statistical analysis. An explanation for ninety-five percent of the variability of the relationship between the independent variable and dependent variable is the percent commonly accepted as having statistical significance.[188] Knowing this provides necessary insight for evaluating claims of 30.2, 67, or even 76 percent correlation between the independent and dependent variables.[189]

This is useful in developing metrics because the complexity of nation assistance operations may prevent the discovery of a direct relationship between the dependent and independent variable. Determining the degree of correlation in the relationship becomes necessary for determining whether the effect between the dependent and independent variable is replicable or happened by chance.[190] Action or decision will invariably rely on when correlation can be paired to experience that shows the correlation to be a plausible relationship.

---

[188] Campbell & Campbell, 2006, 40.

[189] Gorkowski (2009) found 30.2 percent of the variability in increasing violence is explained by CERP expenditure, Clarke and Onufer (2009) found that 67 of the variability in increasing violence is explained by increasing unemployment, O'Connell (2008) found that 76 percent of the variability in Taliban attacks is explained by increases in security projects.

[190] Campbell & Campbell, 2006, 40.

2. What is Not Effective and What is Not Useful in Determining Metrics that Measure Outcomes

a. Overly Complex Qualitative and Overly Detailed Quantitative Metrics that Fail to Measure Outcomes

Iraq and Afghanistan have both seen the use of qualitative and quantitative assessments. Qualitative methods have created metrics too difficult for military units to operationalize. An example of a qualitative method is the TCAPF. Quantitative methods require too many indicators for reporting and focus more on measuring outputs than outcomes. An example of a quantitative method is the MPICE, which has over 800 metrics.[191]

b. Data Analysis

In order to conduct data analysis, a dependent variable must be selected. The problem is immediate as there is no good dependent variable to use for nation assistance. The dependent variable needs to represent the goal of promoting long-term regional stability. Violence against coalition forces, violence against civilians, or number of tips reported have been used as the dependent variable. Unfortunately, due to the complexity of nation assistance and the complexity of the goal of nation assistance, choosing any of the aforementioned examples as a dependent variable may be lacking. Levels of violence may rise or fall independent of stability increasing.

The next problem with data analysis is that the data "is only as accurate as the discipline, reporting standards, and priorities of those units reporting."[192] When analyzing data, it is almost impossible to know the circumstances in which the data was collected. The data set may tell you exactly where, when, and how many improvised explosive devices (IEDs) were found, enemy direct attacks occurred, or tips called-in. However, the data does not tell

---

[191] Dziedzic, Sotirin, Agoglia, 2008.

[192] Beskow notes duplicate reporting, incomplete reports, and reports lacking geographical data as common problems. Beskow, 2011.

you if more friendly or enemy troops moved into the area, if friendly or enemy troops changed their TTPs, or if there was a change in the civilian population. As a result, findings from the data can be scrutinized with questions that cannot be answered.

3. Tactical Outcome Assessment

The Tactical Outcome Assessment is a system designed for tactical units to measure outcomes in their OE. The Tactical Outcome Assessment operationalizes the Failed State Index developed by the Fund for Peace for a tactical unit. The logical lines of operations (LLO) in Army doctrine *FM 3-24 COUNTERINSURGENCY* were used to filter the indicators from the Failed State Index. Changes were also made in restructuring the indicators to read more like measurements, since the goal of the indicators is to measure outcomes.

Below is the Tactical Outcome Assessment with its seven indicators.

1. Increasing/Decreasing Security
2. Increasing/Decreasing Governance
3. Increasing/Decreasing Economic Opportunities
4. Increasing/Decreasing Access to Essential Services
5. Increasing/Decreasing Rule of Law
6. Increasing/Decreasing Movement/Number of Refugees or Internally Displaced Persons
7. Increasing/Decreasing Return/Flight of Influential Individuals

The next step in using the Tactical Outcome Assessment is to determine what metric will represent each individual indicator. This responsibility is up to the tactical unit, and a BCT or JSOTF staff—through mission analysis, intelligence preparation of the battlefield (IPB), and establishing commander's critical information requirements (CCIR)—already has the knowledge base to develop metrics representative of the seven indicators for their OE.

Metrics representative for most of the indicators can be found by simply asking, "What does right look like?" The answers to this question will identify

indicators and provide a basis for which to be measured. If the answers cannot be seen or found, they are not good indicators.

Asking, “What does right look like?” keeps the questions simple. Keeping the questions to yes and no responses with a variable scale from one to ten standardizes the process and allows for aggregating the scores like in the Failed State Index. The scores for each indicator will then be aggregated to determine a single number for their area. Liberty is being exercised in aggregating the numbers. The numbers being aggregated are not cardinal numbers. However, coming up with a single number as a score is beneficial as it aids in showing trends.

B. RECOMMENDATIONS

1. Implementation

The metric reporting requirements to higher headquarters most likely will not track outcomes, or will be too general to be the most effective system to use in a specific area. A grassroots effort should be initiated by BCT and JSOTF staff officers to implement outcome-focused metrics, built from the Tactical Outcome Assessment, into the unit’s PIR. Inputting into the PIR will ensure that resources are diverted to reporting on the outcome-focused metrics.

2. Areas for Further Research

Areas for further research identified include testing the Tactical Outcome Assessment on a nation assistance operation. Testing the Tactical Outcome Assessment will identify whether or not it is effective at actually measuring outcomes as well as identifying areas where it can be improved.

THIS PAGE INTENTIONALLY LEFT BLANK

LIST OF REFERENCES

About the Fund for Peace. 2011. http://www.fundforpeace.org/global/?q=aboutus (accessed August 21, 2011).

Beskow, David. *Spatial and Temporal Analysis of Tactical Integrated Ground Reporting.* Master's thesis, Monterey: Naval Postgraduate School, 2011.

Brooks, David. *Smart Power Setback.* June 20, 2011. http://www.nytimes.com/2011/06/21/opinion/21brooks.html (accessed July 07, 2011).

Brooks, Michael, ed. *Measures of Effectiveness in Stability Operations.* Vols. No. 10-41. Ft. Leavenworth, Kansas: Center for Army Lessons Learned, 2010.

Burke, Mark C., and Eric C. Self. *POPULATION ANALYSIS: A METHODOLOGY FOR UNDERSTANDING POPULATIONS IN COIN ENVIRONMENTS.* December 2008. http://www.dtic.mil/cgi-bin/GetTRDoc?Location=U2&doc=GetTRDoc.pdf&AD=ADA494013 (accessed March 15, 2011).

Campbell, T. Colin, and Thomas M. Campbell II. *The China Study.* Dallas: BenBella Books, Inc, 2006.

Cave, Damian. *Sadr City starts to turn around, posing new challenges - Africa & Middle East - International Herald Tribune.* February 08, 2007. http://www.nytimes.com/2007/02/08/world/africa/08iht-sadr.4525914.html (accessed May 27, 2011).

CIDNE. 2011. http://www.issinc.com/programs/cidne.html (accessed October 11, 2011).

Clarke, Jason A., and Tracy L. Onufer. *UNDERSTANDING ENVIRONMENTAL FACTORS THAT AFFECT VIOLENCE IN SALINAS, CALIFORNIA.* December 2009. http://www.dtic.mil/cgi-bin/GetTRDoc?Location=U2&doc=GetTRDoc.pdf&AD=ADA514256 (accessed October 10, 2011).

Cohen, Craig. "Measuring Progress in Stabilization and Reconstruction [PDF]." United States Institute of Peace Stabilization and Reconstruction, March 2006.

Cohen, Elliot, Conrad Crane, Jan Horvath, and John Nagle. "Principles, Imperatives, and Paradoxes of Counterinsurgency." *Military Review*, March-April 2006: 49–53.

Daddis, Gregoray A. *No Sure Victory: Measuring U.S. Army Effectiveness and Progress in the Vietnam War.* New York: Oxford University Press, 2011.

Dilegge, Dave. *Hearts and Minds.* October 21, 2007. http://smallwarsjournal.com/blog/2007/10/hearts-and-minds/ (accessed March 11, 2011).

Dobbins, James, et al. *America's Role in Nation-Building: From Germany to Iraq.* 2003. http://www.rand.org/pubs/monograph_reports/MR1753.html (accessed August 31, 2011).

Dobbins, James, Seth G. Jones, Crane Keith, and Cole Degrasse Beth. *The Beginner's Guide to Nation Building.* 2007. http://www.rand.org/pubs/monographs/MG557.html (accessed August 31, 2011).

Downes-Martin, Stephen. "Operations Assessment in Afghanistan is Broken: What is to be Done? [PDF]." *Naval War College Review*, Autumn 2011: 103–125.

Dziedzic, Michael, Barabara Sotirin, and John Agoglia. *Measuring Progress in Conflict Environments (MPICE): A Metrics Framework for Assessing Conflict Transformation and Stabilization.* August 2008. http://www.usip.org/files/resources/MPICE%20Aug%2008.pdf (accessed November 21, 2011).

Etzioni, Amitai. *Beware of Generals Carrying Metrics.* September 30, 2010. http://www.huffingtonpost.com/amitai-etzioni/beware-of-generals-carryi_b_745343.html?ref=email_share (accessed June 13, 2011).

Evans, Edward R., and James R. Spies. *INSURGENCY IN THE HOOD: UNDERSTANDING INSURGENCIES THROUGH URBAN GANGS.* Thesis, Defense Analysis, Naval Postgraduate School, Monterey: Naval Postgraduate School, 2006.

Fall, Bernard B. "The Theory and Practice of Insurgency and Counterinsurgency." *The Naval War College Review*, April 1965: 46–56.

Field Manual No. 3-07 STABILITY OPERATIONS. Washington, DC: Headquarters Department of the Army, 2008.

Field Manual No. 3-24 COUNTERINSURGENCY. Washington, DC: Headquarters Department of the Army, 2006.

Fields, Arnold. "Testimony of Arnold Fields Special Inspector General for Afghanistan Reconstruction Before the Commission on Wartime Contracting." *SIGAR.mil.* February 22, 2010. http://www.sigar.mil/pdf/testimony/SIGAR-10-001T.pdf (accessed November 01, 2010).

Frankum, Ronald B., JR., and Stephen F. Maxner. *The Vietnam War Opposition in America Finds Meaning.* http://www.dummies.com/how-to/content/the-vietnam-war-opposition-in-america-finds-meanin.html (accessed March 03, 2011).

Frontline: Interview with Col. Kalev Sepp (Ret.). June 19, 2007. http://www.pbs.org/wgbh/pages/frontline/endgame/interviews/sepp.html (accessed November 03, 2011).

George, Alexander L., and Andrew Bennett. *Case Studies and Theory Development in the Social Sciences.* Cambridge: MIT Press, 2005.

Gordon, Michael R. *War Over Wall Persists in Sadr City Despite Truce.* May 15, 2008. http://www.nytimes.com/2008/05/15/world/middleeast/15wall.html (accessed May 31, 2011).

Gorkowski, Justin B. *A PENNY FOR YOUR THOUGHTS, A NICKEL FOR YOUR HEART: THE INFLUENCE OF THE COMMANDER'S EMERGENCY RESPONSE PROGRAM ON INSURGENCY .* December 2009. http://www.dtic.mil/dtic/tr/fulltext/u2/a514272.pdf (accessed October 10, 2011).

Guttieri, Karen. "Metrics in Iraq's Complex Conflict Environment." In *The Three Circles of War*, by Heather S. Gregg, Hy S. Rothstein and John Arquilla, 137–151. Dulles, VA: Potomac Books, 2010.

Hollinger, Mathew E. *MUQTADA AL-SADR: HOW TO DEMILITARIZE AL-SADR.* Thesis, Monterey: Naval Postgraduate School, 2009.

Hubbard, Andrew P. *A COUNTRY DIVIDED: THE IMPACTS OF FRAGMENTED COMMUNITIES ON IRAQ'S GOVERNMENT.* Thesis, Monterey: Naval Postgraduate School, 2007.

Joint Publication 3-22: Foreign Internal Defense. 2010.

Karnow, Stanley. *Vietnam: A History, Second Edition.* New York: Penguin Books, 1997.

Kilcullen, David. *Counterinsurgency.* New York, NY: Oxford University Press, 2010.

—. *The Accidental Guerrilla.* New York, NY: Oxford University Press, Inc., 2009.

Krepinevich, Andrew F., Jr. *The Army and Vietnam.* John Hopkins University Press, 1986.

Levitt, Steven D., and Dubner J. Stephen. *Freakonomics: A Rogue Economist Explores the Hidden Side of Everything.* New York: Harper Collins, 2006.

—. *Super Freakonomics.* New York: HarperCollins, 2009.

"Logic Model Development Guide [PDF]." W.K. Kellogg Foundation, 2004.

Mansoor, Peter R., Colonel, and Mark S., Major Ulrich. "Linking Doctrine to Action: A New COIN Center-of-Gravity Analysis." *Military Review.* September-October 2007.

McNaught, Anita. *Winning the Peace in Iraq by Rebuilding Sadr City.* June 26, 2008. http://www.foxnews.com/story/0,2933,371764,00.html (accessed May 27, 2011).

Meharg, Sarah Jane. "Measuring the Effectiveness of Reconstruction and Stabilization Activities." *Pearson Papers, Measures of Effectiveness: Peace Operations and Beyond, Volume 10, Issue 1.* Edited by Ann Livingstone. Clementsport, Nova Scotia: The Pearson Peacekeeping Center, 2007.

Middle East: Iraq. May 17, 2011. https://www.cia.gov/library/publications/the-world-factbook/geos/iz.html (accessed May 22, 2011).

Nagl, John A. *Learning to Eat Soup with a Knife: Counterinsurgency Lessons from Malaya and Vietnam.* Chicago: University of Chicago Press, 2005.

O'Connell, Thomas J. *AFGHANISTAN RECONSTRUCTION - A QUANTITATIVE ANALYSIS OF THE INTERNATIONAL EFFORT.* March 2008. http://www.dtic.mil/cgi-bin/GetTRDoc?AD=ADA479990&Location=U2&doc=GetTRDoc.pdf (accessed October 10, 2011).

Powell, Colin. *My American Journey.* New York, New York: Random House, Inc., 1995.

Reconstruction: Brig. Gen. Talley, Mr. Nazar Al-Sultan, Mr. Bass, Jan. 15. January 15, 2009. http://www.usf-iraq.com/?option=com_content&task=view&id=24945&Itemid=131 (accessed May 31, 2011).

Republic of Iraq District Government Field Manual. PDF, USAID, 2007.

Rumsfeld, Donald. *Rumsfeld Memo.* October 13, 2003. http://www.usatoday.com/news/washington/executive/rumsfeld-memo.htm (accessed January 27, 2011).

Sadr City. May 22, 2011. http://topics.nytimes.com/top/news/international/countriesandterritories/iraq/sadr_city/index.html (accessed May 22, 2011).

Schelling, Thomas C. *Arms and Influence.* New Haven, Connecticut: Yale University Press, 1966.

Schroden, Jonathan. "Why Operations Assessments Fail: It's Not Just the Metrics." *Naval War College Review* 64, no. 4 (Autumn 2011): 89–102.

Sepp, Kalev I. "Best Practices in Counterinsurgency." *Military Review*, May/June 2005: 8–12.

Shore, Zachary. *Blunder.* New York: Bloomsbury USA, 2008.

Singer, Peter W. *The Future of National Security, By the Numbers.* May 22, 2011. http://www.brookings.edu/articles/2011/05_national_security_singer.aspx (accessed May 22, 2011).

Sutherland, Sean P. *Measuring Effectiveness in Conflict Environments.* September 2009. http://www.dtic.mil/cgi-bin/GetTRDoc?AD=ADA509874&Location=U2&doc=GetTRDoc.pdf (accessed January 27, 2011).

Thompson, Edwina. *WINNING 'HEARTS AND MINDS' IN AFGHANISTAN: ASSESSING THE EFFECTIVENESS OF DEVELOPMENT AID IN COIN OPERATIONS .* April 2010. http://www.wiltonpark.org.uk/resources/en/pdf/22290903/22291297/wp1022-report (accessed July 23, 2011).

Totten, Michael. *Sadr City After the Fall.* April 8, 2009. http://www.michaeltotten.com/archives/2009/04/sadr-city-after.php (accessed May 25, 2011).

Tschirgi, Glen. *Sadr City's Lesson for Gaza.* January 13, 2009. http://www.americanthinker.com/2009/01/does_sadr_city_have_a_lesson_f.html (accessed May 25, 2011).

US chief Petraeus vows to protect Afghan civilians. July 1, 2010. http://www.bbc.co.uk/news/10472555 (accessed March 11, 2011).

USAID. *Measuring Effectiveness to Improve Effectiveness.* June 2010. www.countrycompass.com/_docs/policy_briefs/Briefing_Note_5_Effectiveness.pdf (accessed February 06, 2011).

Wendt, Eric P. "Strategic Counterinsurgency Modeling." *Special Warfare*, September 2005: 2-13.

Wilson, David, and Gareth E. Conway. "The Tactical Conflict Assessment Framerwork: A Short-lived Panacea." *Informaworld.com.* February 2009. http://www.informaworld.com/smpp/section?content=a909918221&fulltext=713240928 (accessed September 11, 2010).

WITS Methodology and Criteria. https://wits.nctc.gov/FederalDiscoverWITS/in (accessed October 11, 2011).

INITIAL DISTRIBUTION LIST

1. Defense Technical Information Center
   Ft. Belvoir, Virginia

2. Dudley Knox Library
   Naval Postgraduate School
   Monterey, California

3. Hy Rothstein
   Naval Postgraduate School
   Monterey, California

4. Henry James
   Naval Postgraduate School
   Monterey, California

5. Joint Special Operations University
   Hurlburt Field, Florida

6. ASD/SOLIC
   Washington, D.C.

7. U.S. Special Operations Command (USSOCOM) J-7
   MacDill AFB, Florida

8. HQ USSOCOM Library
   MacDill AFB, Florida

9. Commander, 95th Civil Affairs Brigade
   Fort Bragg, North Carolina

10. DCO, 95th Civil Affairs Brigade
    Fort Bragg, North Carolina

11. Daniel Fuhr
    Naval Postgraduate School
    Monterey, California

12. Hieu Pham
    Naval Postgraduate School
    Monterey, California